

Global Investigation of Soil Moisture and Latent Heat Flux Coupling Strength

Fangni Lei¹, Wade T. Crow¹, Thomas R. H. Holmes², Christopher Hain³,
and Martha C. Anderson¹

¹Hydrology and Remote Sensing Laboratory, USDA Agricultural Research Service, Beltsville,
MD 20705, USA

²Hydrological Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD
20771, USA

³Earth Science Office, NASA Marshall Space Flight Center, Huntsville, AL 35805, USA

Corresponding author: Fangni Lei (E-mail: Fangni.Lei@ars.usda.gov)

Key Points:

- Multiple sources of remote sensing-based soil moisture and latent heat flux products are integrated via triple collocation
- Global observation-based estimates of coupling strength between surface soil moisture and latent heat flux are obtained
- Land surface models overestimate the strength of this coupling along transitional climate regimes

Abstract

As a key variable in the climate system, soil moisture (SM) plays a central role in the earth's terrestrial water, energy, and biogeochemical cycles through its coupling with surface latent heat flux (LH). Despite the need to accurately represent SM/LH coupling in earth system models, we currently lack quantitative, observation-based, and unbiased estimates of its strength. Here, we utilize the triple collocation (TC) approach introduced in Crow et al. (2015) to SM and LH products obtained from multiple satellite remote sensing platforms and land surface models (LSMs) to obtain unbiased global maps of SM/LH coupling strength. Results demonstrate that, relative to coupling strength estimates acquired directly from remote sensing-based datasets, the application of TC generally enhances estimates of warm-season SM/LH coupling, especially in the western United States, the Sahel, Central Asia, and Australia. However, relative to triple collocation estimates, LSMs (still) over-predict SM/LH coupling strength along transitional climate regimes between wet and dry climates, such as the central Great Plains of North America, India, and coastal Australia. Specific climate zones with biased relations in LSMs are identified to geographically focus the re-examination of LSM parameterizations. TC-based coupling strength estimates are robust to our choice of LSM contributing SM and LH products to the TC analysis. Given their robustness, TC-based coupling strength estimates can serve as an objective benchmark for investigating model predicted SM/LH coupling.

Plain Language Summary

Physical models describing land-atmosphere coupling have been developed to help better understand the impact of local, regional, and global-scale climate on weather and the water cycle. However, verifying the accuracy of these models is challenging over sparsely instrumented areas. Here, the strength of land-atmosphere coupling between soil moisture and terrestrial evapotranspiration is examined by combining multiple global-scale remote sensing and modeling products into a unified analysis. This analysis is unique in that it can be conducted globally and is unbiased by the presence of random errors in the remote sensing products. As such it provides the first robust estimate of the degree to which soil moisture and evapotranspiration are linked. Results show strong soil moisture/evapotranspiration coupling over the Western United States, the African Sahel, Central Asia, and Australia. However, they also demonstrate that most existing models are still over-predicting this coupling along transitional regions between wet and dry climates (like the Central Great Plains of North America, India, and coastal Australia). This work will help improve the representation of land-atmosphere coupling in models used to obtain future climate projections.

1. Introduction

Soil moisture (SM), through its direct limiting effect on surface latent heat flux (LH), modulates feedbacks between the land surface and the lower atmosphere (Koster et al., 2004; Guo et al., 2006; Jung et al., 2010; Combe et al., 2016). For regions with a strong causal relationship between SM and surface LH, this coupling can significantly affect terrestrial water, energy, and biochemical cycles in a changing climate (Seneviratne et al., 2010, 2013). Despite the widespread application of land surface models (LSMs) to conceptually diagnose this coupling relationship (Van den Hurk et al., 2011), there is no clear consensus on the true strength of terrestrial SM/LH coupling across the globe. Moreover, global observation-based quantification and validation of this relationship within LSMs have not yet been attempted (Dirmeyer et al., 2006a, 2009, 2018; Best et al., 2015).

The limiting effect of SM on LH is generally characterized as a first-order causal relationship within LSMs (Dirmeyer et al., 2006b). Validation practices of a model's ability to characterize this coupling are commonly based on sampling mutual information proxies (e.g., correlation coefficient or the coupling index proposed by Dirmeyer et al., 2011) acquired from collocated SM and LH time series observations (Dirmeyer et al., 2009, 2018; Loew et al., 2017). However, these proxies are usually biased low due to the existence of random measurement errors in any instrument-based observation (i.e., remotely-sensed retrievals or ground measurements) (Crow et al., 2015; Findell et al., 2015) and scale representative differences between local observations and a (coarser-scale) modeling grid (Dirmeyer et al., 2018). Therefore, unbiased, observation-based estimates of SM/LH coupling strength are difficult to acquire.

Recent progress in triple collocation (TC) analysis ([McColl et al., 2014](#); [Gruber et al., 2016a](#)) offers a strategy for deriving unbiased estimates of SM/LH coupling strength. Specifically, if triplets of LH and SM can be obtained from mutually independent sources, TC-based approaches allow us to estimate the variance of random errors affecting all three estimates of both LH and SM. This information, in turn, can be used to compensate sampled estimates of SM/LH coupling strength for the (degrading) impact of these errors on sampled correlations between SM and LH.

Using collocated SM and LH estimates obtained from ground measurements, remote sensing (RS) and a LSM, Crow et al. ([2015](#)) provided a proof-of-concept for the robust estimation of the SM/LH coupling strength via TC analysis. However, Crow et al. ([2015](#)) focused only on a handful of sites across the United States where long-term, high-quality, ground-based SM and LH measurements are available. Moving from their site-based approach to a continuous global investigation requires the simultaneous availability of two satellite-based products for SM and LH that are sufficiently independent for a TC analysis. Fortunately, recent advances in remotely-sensed SM using passive and active microwave sensors ([Kumar et al., 2018](#)) and RS-based LH using satellite-based land surface temperature (LST) products obtained from both thermal-infrared and microwave sensors ([Anderson et al., 2011](#); [Holmes et al., 2018](#)) provide such an opportunity.

Bolstered by this newly developed TC method, and the simultaneous availability of remotely-sensed SM and LH datasets from multiple independent sources, this work seeks to obtain unbiased global estimates of terrestrial SM/LH coupling strength. Section 2 reviews the TC approach previously developed by Crow et al. ([2015](#)) as well as the RS- and LSM-based datasets used to construct the data triplets required by this approach. Results are

presented in Section 3 and discussed in Section 4. Main conclusions are then summarized in Section 5.

2. Method and Datasets

2.1 Triple Collocation-based Coupling Strength Metric

For geophysical variables with at least three independent estimates, the mutually uncorrelated random error variances contained within these estimates can be determined through the triple collocation (TC) method (Stoffelen, 1998; McColl et al., 2014; Gruber et al., 2016a). This approach assumes each of the three estimates (X_i) follows a linear error model:

$$X_i = \alpha_i + \beta_i T + \varepsilon_i, \quad i \in [1, 2, 3] \quad (1)$$

where α and β are temporally-constant additive and multiplicative biases, and ε is zero-mean random error which is independent of the truth T , i.e., error orthogonality. Based on these assumptions, the variances and the covariances for the three datasets can be written as:

$$Var[X_i] = \beta_i^2 Var[T] + Var[\varepsilon_i] \quad (2)$$

$$Cov[X_i, X_j] = \beta_i \beta_j Var[T] + Cov[\varepsilon_i, \varepsilon_j] \quad (3)$$

where $i, j \in [1, 2, 3]$, and the operators $Cov[\cdot]$ and $Var[\cdot]$ denote the covariance and temporal variance, respectively. In TC applications, the three datasets are usually assumed to have mutually independent errors and therefore $Cov[\varepsilon_i, \varepsilon_j] = 0$ when $i \neq j$. Thus, from (2) and (3), the true signal variance (using one dataset as a reference) and random error variances of each dataset can be derived as:

$$Var[T]_i = \beta_i^{-2} \frac{Cov[X_i, X_j] Cov[X_i, X_k]}{Cov[X_j, X_k]} \quad (4)$$

$$Var[\varepsilon_i] = Var[X_i] - \frac{Cov[X_i, X_j]Cov[X_i, X_k]}{Cov[X_j, X_k]} \quad (5)$$

where $i, j, k \in [1, 2, 3]$ and $i \neq j \neq k$. Here, X_i is chosen as the reference to obtain the true signal variance $Var[T]_i$.

Instead of estimating the random error variance of an unknown signal observed by three datasets as described above, Crow et al. (2015) extended TC to estimate the coupling strength between two different geophysical variables, namely SM and LH. The basic assumptions are same as in classic TC, that is, all measurements are linearly related to their corresponding true values

$$SM_A = \alpha_A + \beta_A SM_{True} + \varepsilon_A, \quad A \in [i, j, k] \quad (6)$$

$$LH_B = \alpha_B + \beta_B LH_{True} + \varepsilon_B, \quad B \in [l, m, n] \quad (7)$$

where α and β represent additive and multiplicative systematic errors of each dataset with regard to the (unknown) truth $True$, ε denotes zero-mean random error with temporally constant variance, and A and B represent three independent estimates of SM and LH, respectively. By further assuming the mutual independence of random errors ε within each triplet and the orthogonality of all errors with respect to its respective truth, the coefficient of determination R^2 between two different geophysical variables can be solved as:

$$\begin{aligned} R^2[SM_{True}, LH_{True}]_{i,l} &\equiv \frac{Cov[SM_{True}, LH_{True}]_{i,l}^2}{Var[SM_{True}]_i Var[LH_{True}]_l} \\ &= \frac{Cov[SM_i, LH_l]^2 Cov[SM_j, SM_k] Cov[LH_m, LH_n]}{Cov[SM_i, SM_j] Cov[SM_i, SM_k] Cov[LH_l, LH_m] Cov[LH_l, LH_n]} \end{aligned} \quad (8)$$

where $i \neq j \neq k$ and $l \neq m \neq n$. The advantage of (8), as opposed to the direct sampling of R^2 between any two SM and LH products, is that (8) is unbiased in the presence of random observational errors. Therefore, it provides a means for estimating true coupling strength

between SM and LH using error-prone data (e.g., remote-sensing retrievals of SM_A and LH_B).

A more detailed derivation for (8) can be found in Crow et al. (2015).

Note that an arbitrary “reference pair” of SM_i and LH_l products is applied within the numerator of (8). If all such pairings are free from cross-correlated estimation errors ($Cov[\varepsilon_i, \varepsilon_l] = 0$), the use of any single pair should not bias correlation results. Thus, considering all possible combinations of SM_i and LH_l from each corresponding triplet, nine estimates of $R^2[SM_{True}, LH_{True}]_{i,l}$ can be obtained in total from (8). However, certain reference pairs may possess correlated LH and SM errors, specifically when both SM_i and LH_l are obtained from the same LSM. Therefore, a robust final estimate of $R^2[SM_{True}, LH_{True}]$ can be achieved by taking the median of results obtained from reference pairs deemed to include mutually independent SM and LH errors. The reference pairs used here are fully described below in Section 2.5.

In real-case applications, Crow et al. (2015) acquired their SM and LH triplets from mutually independent ground-, RS- and LSM-based sources. However, this strategy limited the geographic extent of their approach to only a handful of heavily instrumented sites in the United States. Therefore, a key challenge is substituting new remotely-sensed sources of LH and SM estimates into the analysis (to eliminate Crow et al. (2015)’s reliance on ground data) without violating the error independence assumption on which TC is based. More detailed explanations are included in Section 3.1 and the impact of potential violation of assumptions will be discussed in Section 4.1.

2.2 Remotely-sensed Soil Moisture Datasets

Given that the physical mechanism for retrieving SM using active scatterometer versus passive radiometer varies considerably, simultaneous SM retrievals acquired from both active and passive microwave sensors, in combination with LSM estimates, are commonly utilized together in TC applications (Draper et al., 2013; Su et al., 2014; Gruber et al., 2016a).

Here, global passive microwave (PMW) SM retrievals were obtained from the European Space Agency (ESA) Climate Change Initiative (CCI) product version 03.2 (Dorigo et al., 2017; Dorigo et al., 2018) between 1978 and 2015 at 0.25° spatial resolution (<https://www.esa-landcover-cci.org/>). This product was generated using a rigorous blending scheme that combines all Level 2 SM retrievals from available sensors into a (single) weighted optimal estimate using their corresponding error variances (Gruber et al., 2017). All PMW SM retrievals were derived via the same Land Parameter Retrieval Model (LPRM, Owe et al., 2008) that interprets brightness temperature measurements of several sensors. From 2007 onwards, this record includes: the Advanced Microwave Scanning Radiometer-Earth Observing System (AMSR-E) on board the NASA EOS Aqua satellite (Njoku et al., 2003), WindSat on board the Coriolis satellite (Parinussa et al., 2012), the Microwave Imaging Radiometer using Aperture Synthesis (MIRAS) on board Soil Moisture and Ocean Salinity (SMOS) mission (Kerr et al., 2010), and AMSR2 on board Global Change Observation Mission-Water (GCOM-W1) (Parinussa et al., 2015). Considering that more isothermal conditions in near-surface temperature profiles benefit the retrieval of surface soil moisture from passive microwave sensors, only night-time or early-morning retrievals were used in the merged ESA CCI-PASSIVE product (Liu et al., 2012). Note that the quality flags associated with the SM retrievals from each individual sensor were considered to screen for frozen soil

conditions and dense vegetation (where the frequency dependent vegetation optical depth values >0.8 indicate that the canopy is too opaque for interpretation of the soil emission). The merged ESA CCI-PASSIVE SM product has been found to be superior to any single-sensor product (Dorigo et al., 2017).

Scatterometer-based SM retrievals were acquired using the Vienna University of Technology (TU-Wien) change detection algorithm (Naeimi et al., 2009) to C-band (5.255 GHz) Advanced SCATterometer (ASCAT) radar instrument on board the ESA MetOp satellite. The Level 2 SM product can be accessed from the European organization for the exploitation of METeorological SATellites (EUMETSAT) Satellite Application Facility on Support to Operational Hydrology and Water Management (H-SAF) website (<http://hsaf.meteoam.it/soil-moisture.php>). This product has a spatial resolution of 25×25 km with a grid spacing of 12.5 km and was resampled onto a regular 0.25° grid using an inverse-distance-weighted-averaging approach. Active microwave (AMW) SM retrievals from both overpasses at 9:30 a.m./p.m. local solar time were averaged into a daily product from January 2007 onwards. Retrievals were also filtered for frozen soil conditions based on auxiliary flag information.

2.3 Remote Sensing-based Latent Heat Flux Datasets

Diagnostic algorithms are commonly implemented for the derivation of LH estimates from RS data using LST derived from thermal infrared (TIR) imagery. Using the core framework of the two-source energy balance algorithm, LH retrievals can be derived from the Atmosphere Land Exchange Inverse (ALEXI) model using TIR-based LST from both geostationary and polar-orbiting satellite sensors. Here, continuous weekly ALEXI-TIR LH retrievals were derived at a global spatial resolution of 0.05° from 2003 to 2013 using clear-

sky LST retrievals acquired from the Moderate Resolution Imaging Spectroradiometer (MODIS) TIR sensor. For additional details, see Hain and Anderson (2017). Hereinafter, this TIR-based approach will be referred to as the “ALEXI-TIR” algorithm.

LST can also be derived via passive microwave observations — opening up the possibility for deriving a new source of RS-based LH estimates. To this end, Holmes et al. (2015) re-constructed the LST diurnal cycle based on Ka-band brightness temperature observations acquired from a constellation of microwave satellites, including AMSR-E, AMSR-2, and WindSat. This new source of diurnal LST information was then used to run the ALEXI energy balance model to generate MW-based LH estimates at 0.25° global spatial resolution (Holmes et al., 2018). Hereinafter, this MW-based approach will be referred to as the “ALEXI-MW” algorithm.

Meteorological inputs for both the ALEXI-TIR and ALEXI-MW algorithms were acquired from the National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR) and CFS-v2 product (Saha et al., 2010; 2014). Surface albedo and leaf area index were extracted from MODIS MOD43C and MOD15A products, respectively. Note that the most recent global ALEXI-TIR and ALEXI-MW LH products are at a weekly temporal resolution from 2003 to 2013. Recent comparisons of ALEXI-TIR and ALEXI-MW products against in-situ Fluxnet measurements demonstrated satisfactory accuracies in both products (Holmes et al., 2018). Thus, continuous weekly ALEXI-TIR and ALEXI-MW products were used without further masking.

Despite the application of an identical ALEXI modeling framework, LST inputs for ALEXI-TIR and ALEXI-MW (i.e., their key input) are derived from distinct spectral regimes with statistically independent error characteristics (Li et al., 2006). Therefore, our

assumption is that LH estimates derived from different spectral bands will be mutually independent and are therefore suitable for application within the same TC triplet. Additional support for this assumption will be presented below in Section 3.1.

2.4 Land Surface Models

The Global Land Data Assimilation System (GLDAS) is designed to produce global fields of terrestrial water and energy fluxes and storages by incorporating various satellite and ground-based observations into offline simulations of advanced LSMs (Rodell et al., 2004), including Noah version 3.3 (Chen et al., 1996, Ek et al., 2003), Community Land Model (CLM) version 2.0 (Dai et al., 2003), Catchment Land Surface Model (CLSM) F2.5 (Koster et al., 2000), and the Variable Infiltration Capacity (VIC) (Liang et al., 1994) model run in water balance mode. From year 2001 onwards, the meteorological forcing datasets of GLDAS-1 and GLDAS-2 are mainly combined from National Oceanic and Atmosphere Administration (NOAA)/Global Data Assimilation System atmospheric analysis fields, the spatially and temporally disaggregated NOAA Climate Precipitation Center Merged Analysis of Precipitation (CMAP) and Global Precipitation Climatology Project (GPCP) precipitation fields. Additionally, radiation fields are derived using the Air Force Weather Agency's AGRicultural METeorological modeling system (Rodell et al., 2004). GLDAS-2 Noah and CLSM simulations were run at a spatial resolution of 0.25°, and GLDAS-1 CLM and VIC models were at 1°.

Here, each pair of SM and LH fluxes from different LSM simulations was individually averaged from 3-hourly model simulations to daily values. Note that soil layer depths vary across LSMs. In particular, soil moisture estimates from the top first layer were extracted for

Noah, VIC and CLSM using their 0–0.1, 0–0.1 and 0–0.02 m soil layers, respectively. For CLM, the top three soil layers were averaged to a total depth of 0–0.091 m.

2.5 Data Pre- and Post-processing

Before estimating SM/LH coupling strength, both RS- and LSM-based fluxes were pre-processed with consideration of their native spatial, temporal and vertical resolutions (listed in Table 1). First, all RS-based SM and LH retrievals were spatially resampled onto a regular 0.25° grid, and 1.00° LSM-based fluxes were spatially downscaled onto the same 0.25° grid. Second, due to their limited soil penetration capability, remotely-sensed SM retrievals depict a shallower vertical depth support than LSM-based SM estimations, usually on the order of 0 – 5 cm. Here, the exponential filtering approach developed by Albergel et al. (2008) was applied to effectively increase the vertical representation of daily RS-based SM time series by dampening its high-frequency signals. The key time scale parameter required by the filter was assumed to be temporally constant and tuned for each 0.25° grid using the Noah first layer (0–0.1 m) SM product as a comparison dataset. Finally, as mentioned above, the most recent global ALEXI-TIR and ALEXI-MW LH products are weekly summations. As a result, all RS- and LSM-based SM and LH fluxes were averaged from daily to weekly values after the aforementioned processing. Constrained by the temporal coverage of RS-based SM and LH retrievals, our analysis focused on the period 2007 to 2013.

Following Crow et al. (2015), time series of SM and LH products were transformed into temporal ranks before applying (8), which yields coefficients of determination in Spearman rank space. As long as the relationship between SM and LH remains monotonically increasing, these coefficients will not be impacted by nonlinearity in functional dependence

of LH on SM. In addition, given the limited number of weeks (maximum 364) in which all SM and LH data were simultaneously available from January 2007 to December 2013, the coupling strength metric obtained from (8) was applied for all possible combinations of SM/LH reference pairs to reduce sampling uncertainty. In total, eight different estimates of $R^2[SM_{True}, LH_{True}]_{i,l}$ can be acquired: 1) SM_{PMW}/LH_{TIR} , 2) SM_{PMW}/LH_{MW} , 3) SM_{PMW}/LH_{LSM} , 4) SM_{AMW}/LH_{TIR} , 5) SM_{AMW}/LH_{MW} , 6) SM_{AMW}/LH_{LSM} , 7) SM_{LSM}/LH_{TIR} , and 8) SM_{LSM}/LH_{MW} . This list couples all possible SM/LH combinations with the notable exception of SM_{LSM}/LH_{LSM} since error cross-correlation is expected in SM and LH products generated by the same LSM. Subsequently, the median of all eight estimates was utilized as the robust estimate (Huber, 2011) of true SM/LH coupling strength $R^2[SM_{True}, LH_{True}]$. Uncertainty due to limited temporal sampling was quantified via a 500-member boot-strapping analysis (Efron and Tibshirani, 1994). Each median replicate was obtained by sampling half of the original sample size with replacement for all eight reference pairs. The standard deviation of boot-strapped replicates was then used to characterize the sampling error.

Only warm-season SM/LH coupling results were considered in the analysis. This required differential seasonal filtering in various latitude bands which was empirically defined to be comparable with other similar climate studies (Dirmeyer, 2011; Phillips and Klein, 2014; Seneviratne et al., 2013; Teuling et al., 2005). For tropical regions (23.25° N to 23.25° S), all 12 months were included. In the subtropics (23.25° N to 35° N and 23.25° S to 35° S), results were based on either May–October data (in the Northern Hemisphere) or November–April data (in the Southern Hemisphere). In temperate zones (>35° N and >35° S), results were based on June–July–August data (in the Northern Hemisphere) or December–January–February data (in the Southern Hemisphere). Global analyses were conducted for

both original and anomaly time series data. The latter was obtained by subtracting a fixed (2007–2013) seasonal climatology calculated by taking the mean (across the entire 7-year data record) of all SM or LH estimates within a 13-week moving window centered on a particular week of the year.

3. Results

Our global investigation of SM/LH coupling strength is based on weekly RS and LSM time series data aggregated over each 0.25° grid. For each grid, at least 80 collocated samples were required for applying (8). TC generally requires a minimum level of mutual cross-correlation between products in each triplet to yield stable results. Therefore, a minimum Spearman rank coefficient of determination (R^2) of 0.03 [-] was required for all pairwise combinations within both the SM and LH triplets. Pixels not passing this cross-correlation threshold were masked from the analysis. This threshold was determined as a compromise between maintaining sufficient global coverage and filtering non-physical results (Crow et al., 2015). As described above, all global results were based on the period 2007-2013 and the (latitude-dependent) seasonal masking described in Section 2.5.

3.1 Verification at AmeriFlux Tower Sites

The central consideration in applying TC is confidence in the mutual independence of the three estimates selected to comprise the SM and LH triplets. Our two RS-based LH products (ALEXI-TIR and ALEXI-MW) are formulated from the same two-source energy balance model. Therefore, concerns about the mutual independence of their errors are

warranted. Similar cross-dependence concerns exist for RS-based SM products acquired from passive and active microwave approaches.

Therefore, to test the validity of using two different RS products within the same triplet, we replicated the AmeriFlux tower-based results in Crow et al. (2015) via the substitution of ground-based SM and LH observations with comparable RS-based products. That is, the flux tower-based LH is replaced with ALEXI-MW LH estimates and ground-based SM is replaced with SM derived from ASCAT observations. If cross-correlation exists within the two RS-based LH products (ALEXI-MW and ALEXI-TIR) or within the two RS-based SM products (ASCAT and CCI-PASSIVE), these substitutions should be associated with systematic changes in TC-based correlations acquired via (8).

Results from this sensitivity analysis are shown in Figure 1. Note that the black bars plot the same TC estimates as Figure 1 in Crow et al. (2015). By bringing the RS-based ALEXI-MW into the LH triplet (in red), generally consistent TC results in Figure 1 can be obtained with and without the use of flux tower-based LH observations. Likewise, replacing the ground-based SM with ASCAT retrievals (in blue) does not qualitatively change TC results over the AmeriFlux sites. Slightly larger discrepancies are found at the FPE site where TC results derived using ground measurements (black bar) have a higher level of uncertainty than results derived using only model and RS data. This is likely due to the elimination of spatial representative errors (between local-scale ground-based observation and coarse-scale RS- and model-based data) when ground-based observations are not utilized.

This consistency in TC results suggests that, despite their use of the same retrieval algorithm framework, differences in the LST inputs used to force ALEXI-TIR and ALEXI-MW LH are profound enough that their LH retrievals can be considered effectively independent

for TC applications. Therefore, it is appropriate to utilize ALEXI-MW as another source of RS-based LH for the global investigation of SM/LH coupling. Furthermore, the simultaneous use of active and passive-based SM retrievals in the same triplet does not appear to be problematic in the application of (8). Additional discussion on the impact of any residual error cross-correlation existing between RS-based LH and SM products on the TC estimates is provided below in Section 4.1.

3.2 Robustness to Land Surface Model Variations

Another important diagnostic for TC results is robustness to variations in the LSM-based estimates used to construct the data product triplet. Due to variations in the processes and parameterization schemes embedded in LSMs, different models exhibit large discrepancies in the degree of direct SM/LH coupling they predict (see Figure 2). Among the four LSMs considered here, the Noah model (Figure 2a) has the strongest warm-season SM/LH coupling with a global mean and median $R^2[SM_{LSM}LH_{LSM}]$ of 0.53 [-] and 0.70 [-], respectively. On the contrary, SM and LH coupling is relatively weaker in the VIC model which has a global mean $R^2[SM_{LSM}LH_{LSM}]$ of 0.45 [-] and median of 0.54 [-]. Notice that the spatial patterns of SM/LH coupling strength also vary across LSMs. Both Noah (Figure 2a) and VIC (Figure 2b) have shown strong coupling (> 0.5 [-]) between SM and LH over arid regions, such as the Sahara and Central Australia. In contrast, $R^2[SM_{LSM}LH_{LSM}]$ results for CLM (Figure 2c) and CLSM (Figure 2d) suggest moderate to marginal SM/LH coupling in these areas. Differences in the degree of SM/LH coupling strength across LSMs are also observed over western United States, South America, and the Arabian Peninsula. Overall, various LSMs predict distinct magnitudes and spatial patterns in coupling strength.

While these LSM differences are prominent, TC results acquired from (8) should, in theory, be insensitive to our choice of a particular LSM to provide the third member of the data product triplet. To examine the robustness of TC-based SM/LH coupling estimates against various models, SM and LH products from four LSMs have been individually applied in (8) and results are shown in Figure 3. As noted above, grids with low mutual correlation ($R^2[-] < 0.03 [-]$) among either SM or LH triplets are masked. The TC estimates derived with different LSMs dataset yield generally consistent magnitudes of SM/LH coupling strength. In addition, all four cases in Figure 3 illustrate a strikingly similar spatial distribution of areas with enhanced SM/LH coupling. Based on TC estimates, strong coupling between SM and LH are observed over Africa, western North America, the Cerrado region of Brazil, Central Asia, and Australia except for coastal regions with high vegetation density.

Figure 4 shows the variations of direct cross-correlation and TC-based estimates of SM/LH coupling across LSMs. Prior to application of TC, different parameterization schemes in LSMs lead to distinct patterns and magnitudes of coupling strength, and thus large inter-model spread for correlation results ($\sigma < R^2[SM_{LSM}LH_{LSM}] >$) in Figure 4a. However, TC-based coupling strength estimates are generally robust to the choice of LSM and exhibit much lower inter-model spread ($\sigma < R^2[SM_{True}LH_{True}] >$), as shown in Figure 4b. The differences in inter-model spreads are compared along the latitudinal transect in Figure 4c, showing consistency in TC estimates across LSMs. In addition, the temporal sampling errors are estimated through boot-strapping analysis as described in Section 2.5. From Figure 4c, the remaining variability in TC estimates (blue line) has a comparable magnitude with respect to the standard deviation (1-sigma) of estimated sampling error (black line). This suggests that residual LSM-to-LSM variations remaining in TC results can largely be

attributed to inadequate temporal sampling. This lack of bias associated with the use of varying LSMs lends credibility to our interpretation of (8) as an unbiased reflection of true SM/LH coupling strength. Therefore, based on encouraging verification results in Figures 1–4, we will use (8) below to critique levels of SM/LH coupling predicted by LSMs and observed by remotely-sensed SM and RS-based LH products.

3.3 Global Coupling Strength Between Soil Moisture and Latent Heat Flux

SM/LH coupling strength can be directly derived from RS-only datasets (without any TC calculations). Figure 5a shows averaged RS-based estimates ($R^2[SM_{RS}LH_{RS}]$) from SM/LH RS combinations, i.e., CCI-PASSIVE/ALEXI-TIR, CCI-PASSIVE/ALEXI-MW, ASCAT/ALEXI-TIR, ASCAT/ALEXI-MW of remotely-sensed SM and RS-based LH retrievals. Meanwhile, averaged $R^2[SM_{LSM}LH_{LSM}]$ across all four LSMs is shown in Figure 5b. Globally, the $R^2[SM_{RS}LH_{RS}]$ estimates sampled from RS data are generally smaller than 0.4 [-] except for Central Africa. On the contrary, the LSMs predict much higher levels of coupling ($R^2[SM_{LSM}LH_{LSM}] > 0.6 [-]$) over the Africa, the Arabian Peninsula, western Asia, Australia, western North America, and parts of the South America. Comparing RS results in Figure 5a with LSM results in Figure 5b, large discrepancies exist over both arid and transitional regions, where the partition of energy fluxes strongly depends on the soil wetness. Moreover, these significant discrepancies in the estimated magnitudes of SM/LH coupling strength are mostly due to LSM intrinsic coupling being much higher than RS-predicted coupling.

There are two possibilities for explaining this difference. First, LSMs may systematically over-represent the levels of coupling. However, considering the random measurement errors which degrade the accuracy of RS products (Holmes et al., 2018;

Sorensen et al., 2018), a second possibility is that elevated levels of retrieval noise in RS retrievals are causing them to systematically underestimate true levels of coupling (Findell et al., 2015).

As was shown by Crow et al. (2015), TC can be used to distinguish between these two possibilities. In particular, the TC-based approach in (8) corrects for the impact of random errors in observations and provides an unbiased estimate that is unaffected by either noise in the RS retrievals or systematic errors in LSM estimates (see Figure 4). Accordingly, a true coupling estimation can be acquired with the simultaneous availability of SM and LH triplets. Arithmetically averaged TC-based coupling estimates ($R^2[SM_{True}LH_{True}]$) across all four LSMs are shown in Figure 5c. Grey shading reflects the masking of grid cells with insufficient mutual correlations among SM or LH products (see above).

Overall, the LSMs appear to predict somewhat more SM/LH coupling in Figure 5b than TC-based estimates of true coupling strength in Figure 5c. However, there are strong regional variations in these differences and, as shown above in Figure 2 and Figure 4a, strong LSM-to-LSM variations are not reflected in Figure 5b. Therefore, direct comparisons between model predicted and TC estimated coupling strengths are needed to identify regional biases in LSM SM/LH coupling predictions. To this end, Figure 6 plots the global distribution of differences between TC-based ($R^2[SM_{True}LH_{True}]$) and LSM-based ($R^2[SM_{LSM}LH_{LSM}]$) coupling results individually for each LSM. All differences within the 2-sigma confidence interval of $R^2[SM_{True}LH_{True}]$, which has a global averaged standard deviation (obtained via the boot-strapping analysis) of 0.10 [-], have been masked in Figure 6. Therefore, the map indicates regions where LSM SM/LH predictions are significantly different than comparable TC estimates of true coupling strength.

Different LSMs demonstrate distinct spatial patterns of over-/under-coupling relative to our TC baseline. Generally, more over-coupling is found in Noah and CLM than in VIC and CLSM. Comparing Noah and TC-based coupling estimates in Figure 6a, Noah appears to over-predict SM/LH coupling strength (relative to TC) over arid and low-vegetated zones, including the: Sahara Desert, Arabian Desert, Thar Desert in Indian, Kalahari Desert in South Africa, Chihuahuan Desert in North Mexico, and arid areas of Central Australia. In addition, Noah also over-predicts coupling in India and within the Cerrado region of Brazil. Marginal under-couplings are found in northeastern North America and northern Russia.

However, this over-coupling is not universal. In contrast, extensive under-coupling is noted in CLM. Most noticeably, bias in CLM SM/LH coupling appears to have a relationship with underlying vegetation and climate conditions (Figure 6c). Specifically, over-coupling is primarily found in heavily-vegetated humid regions, while under-coupling is seen in arid or semi-arid areas with less SM and sparser vegetation coverage (e.g., the western United States, Central Asia, and Central Australia). Likewise, VIC (Figure 6b) under-couples SM and LH for certain regions with moderate vegetation density (e.g., eastern North America, the Cerrado region of Brazil, and Central Africa). As shown in Figure 6d, CLSM predicts generally similar pattern with VIC except for relatively less SM/LH coupling over the transitional wet/dry climate zone (near 10° N) in Central Africa.

In TC applications, strong seasonality has been shown to produce non-robust TC results (Chen et al., 2016). Therefore, we also re-applied (8) to SM and LH anomaly time series. Anomalies are calculated by subtracting out the long-term seasonal climatology as described in Section 2.5. Results are shown in Figure 7. As expected, without the strong seasonal variations in water availability and insolation, the overall coupling strength of

SM/LH (Figure 7) decreases when compared to Figure 5. However, the LSMs still generally predict more coupling than RS-based estimates. Moreover, the TC-based (anomaly) coupling strength estimates (Figure 7c) are generally much closer to LSM-based predictions (Figure 7b) than RS-based results. Regions with strong SM/LH coupling are distributed mainly over transitional zones and semi-arid regions where LSMs also predict similar spatial patterns.

Comparing the baseline TC-based and various LSMs predicted estimates of SM/LH coupling strength, similar spatial tendencies (with regard to over- versus under-coupling in LSMs) are found for both the anomaly (Figure 8) and original time series (Figure 6). For the anomaly analysis, Noah, CLM, and CLSM generally overestimate the coupling across the globe, among which CLM has the most prominent over-coupled relations. On the other hand, both VIC and CLM underestimate the coupling over western United States and Central Australia. In eastern China, all four LSMs systematically predict less SM/LH coupling. Therefore, taken as a whole, the original (Figure 6) and anomaly (Figure 8) time series analyses demonstrate consistent results with regard to LSM SM/LH coupling bias. Percentages of pixels with over-coupling and under-coupling in various LSMs are summarized for six continents in Figure 9.

3.4 Climate Zone Analysis

It is also potentially useful to examine the performance of LSMs as a function of climate zone. To this end, LSM- and TC-based coupling estimates from both the original and anomaly time series are grouped based on their Köppen-Geiger climate zone (Kottek et al., 2006) classification. Detailed descriptions of each climate zone index are included in Table 2. Mean values of $R^2[-]$ for each climate category are then sorted from high to low mean

annual air temperature in Figure 10. Climate zones containing less than 100 grid cells (in which TC estimates are available) are omitted.

Although LSM- and TC-based coupling estimates differ in exact absolute values, the general trends with regard to different climate zones are consistent. All LSMs predict relatively strong coupling between SM and LH over tropical (As and Aw) and arid (BWk, BWk, BSk, and BSh) climate zones in Figure 10a. For warm temperate climate zones, the coupling of SM/LH significantly increases with the reduced soil wetness from fully humid (Cf) to summer dry (Cs) and winter dry (Cw) conditions. This tendency also exists for colder continental climate zones from Df to Ds and Dw, except for Dsc where summer is usually dry (limited soil water availability) and cool (low temperature). For climate zones with comparable precipitation conditions, SM/LH coupling generally decreases with reduced summer maximum air temperature from Dsa (hot summer) to Dsb(warm summer) and Dsc (cold summer). Moving from the original to anomaly time series (Figure 10b), one exception is found for tropical savannah climate zone where LSMs predict less anomaly coupling.

When compared against the robust TC-based results (in black circles), LSMs slightly overestimate the coupling over tropical arid and semi-arid climate zones, including As (equatorial savannah with dry summer), Aw (equatorial savannah with dry winter), BW (cold arid desert), and BS (hot arid desert). Noah, VIC, and CLSM predict a relatively accurate coupling for regions with adequate precipitation, such as Cf and Df. By contrast, CLM generally predicts more coupling than TC analysis for these climate zones. On the other hand, VIC is the only one (out of four LSMs) to under-represent the coupling over Ds (continental climate with dry summer). In addition, LSMs are systematically over-coupled for Dw (continental dry winter) climate zones.

4. Discussion

Over the past two decades, various ground-based observational networks of land surface fluxes and meteorological states have been established for the validation of RS-based retrievals and model-based simulations (Dorigo et al., 2011; Baldocchi et al., 2001). However, these networks are generally located in either North America or Europe and therefore capture a relatively limited range of land surface and climatic conditions for global investigations. In addition, the spatial scale differences between point-scale ground measurement and grid-averaged RS product or model simulation can imperil the accuracy of coupling estimates acquired from point-scale ground observations. Note that the spatial representative mismatch across measurements can be partially alleviated by using various remotely-sensed products with comparable spatial resolutions for global LSMs. Nevertheless, random measurement errors can still induce bias into the sampled correlation between observed variables (Crow et al., 2015). Recently, using an aggregated surface flux dataset FLUXNET2015, Dirmeyer et al. (2018) explored both the atmospheric leg (sensible heat flux/lifting condensation level) and terrestrial leg (soil moisture/latent heat flux) of land-atmosphere coupling across offline/coupled LSMs and multiple reanalysis products. Their results suggested a potentially over-coupled relationship between SM and LH across LSMs. However, their analysis was confined by both the limited geographical extent of the FLUXNET sites and the strong scale contrast between these point-scale observations and coarse-scale LSM grids. Therefore, an unbiased (and spatially extensive) estimate of the true coupling of two geophysical variables is needed to eliminate both spatial scale differences and compensate for the biasing impact of random measurement errors.

RS-based products are usually subject to random errors originated from instrument sensitivity, sensor calibration, retrieval physics and parameterization (Congalton, 1991). Traditional pairwise correlation calculations that ignore random errors may not be appropriate for characterizing the relationship between two variables. For example, the direct use of RS-based estimates substantially under-predicts SM/LH coupling strength (Figure 5a and 7a). A naïve interpretation of this results would suggest that LSMs strongly over-couple SM and LH over almost all regions of the world (see Figure 5b and 7b). However, when explicitly accounting for the random measurement errors via a TC analysis, a significantly larger value of true SM/LH coupling (Figure 5c and 7c) is obtained as compared to underlying RS-based estimates (Figure 5a and 7a).

Consequently, the diagnosis of LSM coupling biases becomes much more complex and geographically variable. Among the four LSMs considered here, VIC turns out to be generally accurate in describing the relationship between SM and LH (Figure 6b and 8b). Nevertheless, certain regional biases of both over- and under-coupling are still observed in LSMs with respect to TC results (see Figure 6 and Figure 8) due to imperfect model physics. In summary, it is important to consider the impact of random errors in RS products when applying canonical statistical measures (e.g. correlation coefficient, root-mean-square-error) to evaluate or benchmark LSMs. The application of (8) is recommended for related applications, especially when multiple RS products are available.

4.1 Impact of Cross-correlated Error

A critical issue when applying the TC approach in (8) is the potential for mutual error dependency among data members in each triplet, specifically for the LH triplet containing

both ALEXI-TIR and ALEXI-MW products and the SM triplet containing both AMW and PMW retrievals. Previous studies indeed suggest that error cross-correlation also exists for AMW and PMW SM retrievals (Gruber et al., 2016b; Chen et al., 2016). Our analysis of cross-correlation at AmeriFlux tower sites indicates that differences in LST inputs for ALEXI-TIR and ALEXI-MW LH products or retrieval processes (and spectral regimes) for AMW and PMW SM products appear to be sufficient for ensuring error independence and the confident application of TC. However, moving from a handful of sites to global analysis, the existence of residual non-zero error cross-correlation among LH or SM products may not be negligible and its possible effect on estimates of coupling strength needs to be discussed.

Appendix A provides a detailed description of the impact of potentially neglected error cross-correlation on TC-based coupling estimates derived via (8). Analytical derivations show that positive error cross-correlation between two datasets in the triplet would negatively bias the estimated true signal variance when using the third independent member as the reference. Thus, in the case of cross-correlated ALEXI-TIR and ALEXI-MW LH errors, a positively-biased estimate of SM/LH coupling strength would be obtained by utilizing a LSM-based LH product as a reference. In contrast, if either ALEXI-TIR or ALEXI-MW LH is used as the reference, the derived estimate can be negatively-biased. Therefore, cross-correlated errors in RS-based LH products would increase the spread in coupling strength estimates acquired across all possible reference pairs. A similar effect would accompany the presence of correlated errors in ASCAT (AMW) and CCI-PASSIVE (PMW) SM products (see Appendix A for details).

As described in Section 2.5, the sensitivity of our final TC-based coupling results to this spread is minimized here through our use of a median operator to summarize a single

result from across all eight possible reference pair choices. Specifically, four out of the eight reference pairs are based on either LSM-based LH or SM products (as described in Section 2.5) and can potentially lead to biased higher estimates of SM/LH coupling strength in (8). On the contrary, the other four (biased lower) estimates are obtained using RS-based products as references (see Appendix A). This suggests that our approach should be reasonably robust to the presence of error cross-correlation in either RS-based LH and SM products. Note that this interpretation is supported by earlier results shown in Figure 1.

4.2 Implications for Land-atmosphere Coupling

Under all circumstances, our results consistently indicate that the random observational errors in RS-based products greatly complicate their application for benchmarking LSM coupling results. However, by using the robust TC estimates derived in part from RS products (Figure 5c and Figure 7c), regions with over- or under-coupled relationship across different LSMs can be identified (Figure 6 and Figure 8). The spatial extent and distribution of over-coupled predictions are comparably similar among various LSMs over the globe, especially for Noah, CLM, and CLSM. Interestingly, all four LSMs tend to over-predict SM/LH coupling over the central Great Plains of North America, the Sahel, India, and coastal Australia (Figure 6 and 8). These regions are commonly known as “hot-spots” of strong land-atmosphere coupling (Koster et al., 2004). Benchmarking the performance of LSMs over these hot-spot regions is of vital importance for seasonal weather and climate forecasting. On the other hand, all four LSMs tend to slightly under-represent the SM/LH coupling relationship in China. VIC and CLM generally under-predict the coupling over less

intense hot-spots regions, including western North America, Central Asia, and Central Australia.

When coupled with climate models, bias in the LSM's SM/LH coupling strength can affect the surface partitioning of sensible and latent heat fluxes and thus water and energy interactions along the land-atmosphere interface. This can result in biased predictions of extreme event frequencies, such as heatwaves and droughts (Ukkola et al., 2016a, 2016b). Note that different LSMs vary in process representation of soil surface-subsurface hydrology, groundwater storage, and plant water components (Egea et al., 2011). It is therefore critical to pinpoint the exact processes that contribute to the unrealistic characterization of water and energy exchanges in LSMs. However, this is beyond the scope of this work and will be addressed in our forthcoming work by exploring various processes and parameterization schemes in the modular Noah-MP model (Niu et al., 2011).

5. Conclusions

Based on the simultaneous availability of multiple SM and LH products from satellite and LSMs, this work provides an unbiased estimate of SM/LH coupling strength at the global scale. In particular, the TC strategy of Crow et al. (2015) is applied globally, for the first time, to compensate for the impact of random observational errors on sampled correlation statistics — an effect which is commonly neglected. This allows us to obtain quasi-global maps of unbiased SM/LH coupling strength estimates for comparisons against comparable LSM predictions.

Using the TC-based robust estimate, several important conclusions can be drawn. First, it is found that with the existence of random measurement errors, RS-based coupling

strength estimates are prone to large negative biases and thus insufficient for directly benchmarking the true SM/LH coupling relationship. Past studies using ground-based measurements are likely to subject to the same limitations ([Dirmeyer et al., 2018](#); [Ukkola et al., 2016b](#)). By contrast, the TC-based approach in (8) provides a means for compensating for this effect. When observational errors are properly accounted for, significantly higher levels of true SM/LH coupling strength relative to RS-based estimates (Figure 5 and 7) are obtained. Regions with strong SM/LH coupling are found over western Northern America, the Sahel, Central Asia and Australia. Relative to TC-based results, regional areas of over- or under-coupling in LSMs are identified.

Overall, results are consistent with [Dirmeyer et al. \(2018\)](#) in indicating that LSMs generally predict over-coupling. However, large regional and LSM-to-LSM variations are also observed. For example, CLM appears to be under-coupled in semi-arid areas and over-coupled in heavily-vegetated humid regions. In addition, areas of over-coupling appear to be concentrated in regions of globe with strong land-atmosphere coupling. Such identification of regional biases can guide subsequent efforts to improve LSM physics and/or parameterizations. Note that TC-based coupling strength estimates are generally robust to the choice of a particular LSM to provide the third (SM or LH) member of each triplet. Moreover, both the original and anomaly time series analyses yield overall consistent results, lending credibility to our conclusions.

One obvious limitation here is our focus on a set of relatively dated LSM versions provided by the GLDAS-1 project. This is particularly for CLM which has evolved from CLM 2.0 (applied in GLDAS-1 and thus examined here) to CLM 4.5 with some fundamental changes. Future work with more recent LSM versions is clearly needed. Likewise, a daily-

scale SM/LH coupling analysis would be preferable relative to the weekly time-scale used here. Expected enhancements to ALEXI LH products should make this possible in the near future. Improvement to ALEXI LH products are also expected via the inclusion of Clouds and Earth's Radiant Energy Systems (CERES) meteorological forcing datasets (Wielicki et al., 1996; <https://ceres.larc.nasa.gov/>).

In contrast to previous research focusing on the use of ground-based measurements ([Ukkola et al., 2016b](#); [Dirmeyer et al., 2006a; 2018](#)) or multi-model cross-comparison ([Best et al., 2015](#); [Haughton et al., 2016](#)), our results are primarily constructed from multiple independent RS-based observations and model products (free of spatial representativeness differences) and yield an objective benchmark for LSM coupling strength estimates. More importantly, this analysis provides quasi-global results that are not restricted to sparse flux tower locations. As a result, climate zones with either over- or under-coupled SM/LH relations in LSMs are readily identified and can be prioritized for future LSM improvement and multi-model comparison experiments, such as the Land Surface, Snow and Soil moisture Model Intercomparison Project (LS3MIP; [Van den Hurk et al., 2016](#)) conducted as part of the sixth phase of the Coupled Model Intercomparison Project (CMIP6; [Eyring et al., 2016](#)). Moreover, the TC-based coupling metric can also be utilized for analyzing other coupled geophysical variables when its prerequisites are satisfied. For example, the relationship between SM and subsequent (or antecedent) precipitation ([Guillod et al., 2015](#); [Tuttle and Salvucci, 2016](#)) or SM and maximum summertime air temperature ([Miralles et al., 2012](#)).

Acknowledgments

This work was supported by the grant 17-TASNPP17-0095 funded by the NASA Science of Terra, Aqua, and Suomi NPP program. The used GLDAS data are available from the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) (<https://hydro1.gesdisc.eosdis.nasa.gov/data/GLDAS/>). The ESA CCI PASSIVE soil moisture product is from <http://dx.doi.org/10.5285/d2eea061026240eb8a2f9cc64a691338>, and ASCAT data are available from EUMETSAT's H-SAF data portal (<http://hsaf.meteoam.it/soil-moisture.php>). In-situ AmeriFlux data can be obtained from <http://ameriflux.lbl.gov/>. The global ALEXI-TIR and ALEXI-MW LH data products are available upon request.

Appendix A

The accuracy of any TC-based estimate depends strongly on the validity of the error independence assumption. In the presence of non-zero error cross-correlation (Gruber et al., 2016b), direct usage of (4) and (8) can lead to biased estimates. Specifically, if there only exists positive error cross-correlation between dataset X_j and X_k ($Cov[\varepsilon_j, \varepsilon_k] > 0$, $Cov[\varepsilon_i, \varepsilon_j] = 0$, and $Cov[\varepsilon_i, \varepsilon_k] = 0$), the estimated true signal variance $Var[T]_i$ using dataset X_i as the reference can be biased low with regard to actual variance of true signal $Var[T]$

$$Var[T]_i = \beta_i^{-2} \frac{Cov[X_i, X_j]Cov[X_i, X_k]}{Cov[X_j, X_k]} = \frac{Var[T]^2}{Var[T] + \beta_j^{-1}\beta_k^{-1}Cov[\varepsilon_j, \varepsilon_k]} \quad (A1)$$

Instead, when using dataset X_j or X_k as the reference, the estimated true signal variance $Var[T]_j$ or $Var[T]_k$ are positively biased

$$Var[T]_j = Var[T] + \beta_j^{-1}\beta_k^{-1}Cov[\varepsilon_j, \varepsilon_k] \quad (A2)$$

Therefore, the estimated coefficient of determination $R^2[SM_{True}, LH_{True}]_{i,l}$ derived from (8) becomes biased high when SM dataset SM_i and/or LH dataset LH_l are chosen as the reference

$$R^2[SM_{True}, LH_{True}]_{i,l} \equiv \frac{Cov[SM_{True}, LH_{True}]_{i,l}^2}{Var[SM_{True}]_i Var[LH_{True}]_l}$$

$$= \frac{Cov[SM_{True}, LH_{True}]_{i,l}^2 \times \{Var[SM_{True}] + \beta_j^{-1} \beta_k^{-1} Cov[\varepsilon_j, \varepsilon_k]\} \times \{Var[LH_{True}] + \beta_m^{-1} \beta_n^{-1} Cov[\varepsilon_m, \varepsilon_n]\}}{Var[SM_{True}]^2 \times Var[LH_{True}]^2} \quad (A3)$$

where $Cov[\varepsilon_j, \varepsilon_k]$ and/or $Cov[\varepsilon_m, \varepsilon_n]$ are (positive) non-zero cross-correlated errors between SM_j and SM_k and/or LH_m and LH_n . By contrast, in the case of SM_j/LH_m is chosen as the reference pair, the estimated coefficient of determination $R^2[SM_{True}, LH_{True}]_{j,m}$ will be biased low

$$R^2[SM_{True}, LH_{True}]_{j,m} \equiv \frac{Cov[SM_{True}, LH_{True}]_{j,m}^2}{Var[SM_{True}]_j Var[LH_{True}]_m}$$

$$= \frac{Cov[SM_{True}, LH_{True}]_{j,m}^2}{\{Var[SM_{True}] + \beta_j^{-1} \beta_k^{-1} Cov[\varepsilon_j, \varepsilon_k]\} \times \{Var[LH_{True}] + \beta_m^{-1} \beta_n^{-1} Cov[\varepsilon_m, \varepsilon_n]\}} \quad (A4)$$

Analogous biases also exist for the reference pairs of SM_j/LH_n , SM_k/LH_m , and SM_k/LH_n .

Here, eight reference pairs are utilized to obtain estimates of SM/LH coupling strength using (8). Concerning the potential positive cross-correlated errors between ALEXI-TIR and ALEXI-MW LH products and/or ASCAT and CCI-PASSIVE SM products, four reference pairs —using LSM SM or LH as a reference— tend to produce positively biased estimates (A3), i.e., SM_{PMW}/LH_{LSM} , SM_{AMW}/LH_{LSM} , SM_{LSM}/LH_{TIR} , SM_{LSM}/LH_{MW} . Instead, the other four estimates are biased lower (A4) using reference pairs of SM_{PMW}/LH_{TIR} , SM_{PMW}/LH_{MW} , SM_{AMW}/LH_{TIR} , and SM_{AMW}/LH_{MW} . Therefore, the existence of error cross-correlation will tend to increase the spread of coupling strength estimates acquired from eight reference pairs. Nevertheless, given that it is obtained via a median operator (applied

to all eight possible reference choices), final TC-based coupling estimates should be relatively robust to this tendency.

Appendix B

Table B1. Basic attributes of a set of AmeriFlux sites used here and Crow et al. (2015).

| Abbreviation | AmeriFlux Site | Latitude/Longitude |
|--------------|------------------------|--------------------|
| SRM | Santa Rita Mesquite | 31.821°/-110.866° |
| SRC | Santa Rita Creosote | 31.908°/-110.840° |
| WHS | Lucky Hills | 31.744°/-110.052° |
| WKG | Kendall Grasslands | 31.737°/-109.942° |
| TON | Tonzi Ranch | 38.432°/-120.966° |
| AUD | Audubon Grasslands | 31.591°/-110.509° |
| SDH | Sand Hills Dry Valley | 42.069°/-101.407° |
| SUH | Sand Hills Upland Dune | 42.066°/-101.367° |
| ARM | ARM-CART | 36.606°/-97.489° |
| BLO | Blodgett Forest | 38.895°/-120.633° |
| DK1 | Duke Open Field | 35.971°/-79.093° |
| DK2 | Duke Hardwoods | 35.974°/-79.100° |
| DK3 | Duke Pine | 35.978°/-79.094° |
| FPE | Fort Peck | 48.308°/-105.102° |
| NE2 | Mead Irrigated | 41.164°/-96.470° |
| NE3 | Mead Rainfed | 41.180°/-96.440° |
| IB1 | Fermi Agricultural | 41.859°/-88.223° |
| IB2 | Fermi Prairie | 41.841°/-88.241° |

References

Albergel, C., C. Rüdiger, T. Pellarin, J. C. Calvet, N. Fritz, F. Froissard, D. Suquia, A. Petitpa, B. Piguet, and E. Martin (2008), From near-surface to root-zone soil moisture using an exponential filter: an

725 assessment of the method based on in-situ observations and model simulations, *Hydrol. Earth Syst.*
726 *Sci.*, 12, 1323-1337, doi:10.5194/hess-12-1323-2008.

727 Anderson, M. C., W. P. Kustas, J. M. Norman, C. R. Hain, J. R. Mecikalski, L. Schultz, M. P. González-
728 Dugo, C. Cammalleri, G. d'Urso, A. Pimstein, and F. Gao (2011), Mapping daily evapotranspiration at
729 field to global scales using geostationary and polar orbiting satellite imagery, *Hydrol. Earth Syst. Sci.*,
730 15, 223-239, doi:10.5194/hess-15-223-2011.

731 Baldocchi, D., E. Falge, L. Gu, R. Olson, D. Hollinger, S. Running, P. Anthoni, C. Bernhofer, K. Davis, R.
732 Evans, J. Fuentes, A. Goldstein, G. Katul, B. Law, X. Lee, Y. Malhi, T. Meyers, W. Munger, W. Oechel, K.
733 T. Paw U, K. Pilegaard, H. P. Schmid, R. Valentini, S. Verma, T. Vesala, K. Wilson, and S. Wofsy (2001),
734 FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon
735 dioxide, water vapor, and energy flux densities, *Bull. Amer. Meteor. Soc.*, 82, 2415-2434,
736 doi:10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2.

737 Best, M. J., G. Abramowitz, H. Johnson, A. Pitman, G. Balsamo, A. Boone, M. Cuntz, B. Decharme, P.
738 Dirmeyer, J. Dong, M. Ek, Z. Guo, V. Haverd, B. J. van den Hurk, G. S. Nearing, B. Pak, C. Peters-Lidard,
739 J. A. Santanello, L. Stevens, and N. Vuichard (2015), The plumbing of land surface models:
740 benchmarking model performance, *J. Hydrometeor.*, 16, 1425-1442, doi:10.1175/JHM-D-14-0158.1.

741 Chen, F., K. Mitchell, J. Schaake, Y. K. Xue, H. L. Pan, V. Koren, Q. Y. Duan, M. Ek, and A. Betts (1996),
742 Modeling of land surface evaporation by four schemes and comparison with FIFE observations, *J.*
743 *Geophys. Res.*, 101(D3), 7251-7268, doi:10.1029/95JD02165.

744 Chen, F., W. T. Crow, A. Colliander M. H. Cosh, T. J. Jackson, R. Bindlish, R. H. Reichle, S. K. Chan, D. D.
745 Bosch, P. J. Starks, and D. C. Goodrich (2016), Application of triple collocation in ground-based
746 validation of Soil Moisture Active/Passive (SMAP) level 2 data products, *IEEE J. Sel. Topics Appl.*
747 *Earth Observ. Remote Sens.*, 99, 1-14, doi:10.1109/JSTARS.2016.2569998.

748 Combe, M., J. V.-G. de Arellano, H. G. Ouwersloot, and W. Peters (2016), Plant water-stress
749 parameterization determines the strength of land-atmosphere coupling, *Agric. For. Meteorol.*, 217,
750 61-73, doi:10.1016/j.agrformet.2015.11.006

751 Congalton, R. G. (1991), Remote sensing and geographic information system data integration: error
752 sources and research issues, *Photogramm. Eng. Remote Sensing*, 57(6), 677-687.

753 Crow, W. T., F. Lei, C. Hain, M. C. Anderson, R. L. Scott, D. Billesbach, and T. Arkebauer (2015),
754 Robust estimates of soil moisture and latent heat flux coupling strength obtained from triple
755 collocation, *Geophys. Res. Lett.*, 42(20), 8415-8423, doi:10.1002/2015GL065929.

756 Dai, Y., X. Zeng, R. E. Dickinson, I. Baker, G. B. Bonan, M. G. Bosilovich, A. S. Denning, P. A. Dirmeyer,
757 P. R. Houser, G. Niu, and K.W. Oleson (2003), The common land model, *Bull. Am. Meteorol. Soc.*, 84,
758 1013-1023, doi:10.1175/BAMS-84-8-1013.

759 Dirmeyer, P. (2011), The terrestrial segment of soil moisture-climate coupling, *Geophys. Res. Lett.*,
760 38, L16702, doi: 10.1029/2011GL048268.

761 Dirmeyer, P. A., R. D. Koster, and Z. Guo (2006a), Do global models properly represent the feedback
762 between land and atmosphere?, *J. Hydrometeor.*, 7(6), 1177-1198, doi:10.1175/JHM532.1.

763 Dirmeyer, P. A., X. Gao, M. Zhao, Z. Guo, T. Oki, and N. Hanasaki (2006b), GSWP-2: Multimodel
764 analysis and implications for our perception of the land surface, *Bull. Amer. Meteor. Soc.*, 87(10),
765 1381-1398, doi:10.1175/BAMS-87-10-1381.

766 Dirmeyer, P. A., C. A. Schlosser, and K. L. Brubaker (2009), Precipitation, recycling, and land
767 memory: An integrated analysis, *J. Hydrometeor.*, 10(1), 278-288, doi:10.1175/2008JHM1016.1.

768 Dirmeyer, P. A., L. Chen, J. Wu, C.-S. Shin, B. Huang, B. A. Cash, M. G. Bosilovich, S. Mahanama, R. D.
769 Koster, J. A. Santanello, M. B. Ek, G. Balsamo, E. Dutra, and D. M. Lawrence (2018), Verification of
770 Land–Atmosphere Coupling in Forecast Models, Reanalyses, and Land Surface Models Using Flux
771 Site Observations, *J. Hydrometeor.*, 19(2), 375-392, doi:10.1175/JHM-D-17-0152.1.

772 Dorigo, W. A., W. Wagner, R. Hohensinn, S. Hahn, C. Paulik, A. Xaver, A. Gruber, M. Drusch, S.
773 Mecklenburg, P. van Oevelen, A. Robock, and T. Jackson (2011), The International Soil Moisture
774 Network: a data hosting facility for global in situ soil moisture measurements, *Hydrol. Earth Syst.*
775 *Sci.*, 15(5), 1675-1698, doi:10.5194/hess-15-1675-2011.

776 Dorigo, W.; W. Wagner, A. Gruber, T. Scanlon, S. Hahn, R. Kidd, C. Paulik, C. Reimer, R. van der
777 Schalie, and R. de Jeu (2018), ESA Soil Moisture Climate Change Initiative (Soil_Moisture_cci):
778 Version 03.2 data collection. Centre for Environmental Data Analysis, 29 June 2018,
779 doi:10.5285/d2eea061026240eb8a2f9cc64a691338.

780 Dorigo, W., W. Wagner, C. Albergel, F. Albrecht, G. Balsamo, L. Brocca, D. Chung, M. Ertl, M. Forkel, A.
781 Gruber, E. Haas, P. D. Hamer, M. Hirschi, J. Ikonen, R. de Jeu, R. Kidd, W. Lahoz, Y. Y. Liu, and P.
782 Lecomte (2017), ESA CCI Soil Moisture for improved Earth system understanding: state-of-the art
783 and future directions, *Remote Sens. Environ.*, 203, 185-215, doi:10.1016/j.rse.2017.07.001.

784 Draper, C., R. Reichle, R. de Jeu, V. Naeimi, R. Parinussa, and W. Wagner (2013), Estimating root
785 mean square errors in remotely sensed soil moisture over continental scale domains, *Remote Sens.*
786 *Environ.*, 137, 288-298, doi: 10.1016/j.rse.2013.06.013.

787 Efron, B and R. J. Tibshirani (1994), The bootstrap estimate of standard error. In *An Introduction to*
788 *the Bootstrap*, 1st ed.; Chapman and Hall/CRC: Boca Raton, FL, USA; pp. 45-49.

789 Egea, G., A. Verhoef, and P. L. Vidale (2011), Towards an improved and more flexible representation
790 of water stress in coupled photosynthesis–stomatal conductance models, *Agric. For. Meteorol.*,
791 151(10), 1370-1384, doi:10.1016/j.agrformet.2011.05.019.

792 Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley (2003),
793 Implementation of Noah land surface model advances in the National Centers for Environmental
794 Prediction operational mesoscale Eta model, *J. Geophys. Res.*, 108(D22), 8851,
795 doi:10.1029/2002JD003296.

796 Eyering, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor (2016),
797 Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and
798 organization, *Geosci Model Dev.*, 9, 1937-1958, doi: 10.5194/gmd-9-1937-2016.

799 Findell, K. L., P. Gentile, B. R. Lintner, and B. P. Guillod (2015), Data length requirements for
800 observational estimates of land-atmosphere coupling strength, *J. Hydrometeor.*, 16(4), 1615-1635,
801 doi:10.1175/JHM-D-14-0131.1.

802 Gruber, A., W. A. Dorigo, W. Crow, and W. Wagner (2017), Triple Collocation-Based Merging of
803 Satellite Soil Moisture Retrievals, *IEEE Trans. Geosci. Remote Sens.*, 55(12), 6780-6792,
804 doi:10.1109/TGRS.2017.2734070.

805 Gruber, A., C.-H. Su, S. Zwieback, W. Crow, W. Dorigo, and W. Wagner (2016a), Recent advances in
806 (soil moisture) triple collocation analysis, *Int. J. Appl. Earth Obs. Geoinf.*, 45, 200-211,
807 doi:10.1016/j.jag.2015.09.002.

808 Gruber, A., C. H. Su, W. Crow, S. Zwieback, W. Dorigo, and W. Wagner (2016b), Estimating error
809 cross-correlations in soil moisture data sets using extended collocation analysis, *J. Geophys. Res:*
810 *Atmos.*, 121(3), 1208-1219, doi:10.1002/2015JD024027.

811 Guillod, B. P., B. Orlowsky, D. G. Miralles, A. J. Teuling, and S. I. Seneviratne (2015), Reconciling
812 spatial and temporal soil moisture effects on afternoon rainfall, *Nat. Commun.*, 6, 6443,
813 doi:10.1038/ncomms7443.

814 Guo, Z., P. A. Dirmeyer, R. D. Koster, Y. C. Sud, G. Bonan, K. W. Oleson, E. Chan, D. Versegghy, P. Cox, C.
815 T. Gordon, J. L. McGregor, S. Kanae, E. Kowalczyk, D. Lawrence, P. Liu, D. Mocko, C. Lu, K. Mitchell, S.
816 Malyshev, B. McAvaney, T. Oki, T. Yamada, A. Pitman, C. M. Taylor, R. Vasic, and Y. Xue (2006),
817 GLACE: the global land-atmosphere coupling experiment. Part II: analysis, *J. Hydrometeor.*, 7(4),
818 611-625, doi:10.1175/JHM511.1.

819 Miralles, D. G., M. J. van den Berg, A. J. Teuling, and R. A. M. de Jeu (2012), Soil moisture-temperature
820 coupling: A multiscale observational analysis, *Geophys. Res. Lett.*, 39(21), L21707,
821 doi:10.1029/2012GL053703.

822 Hain, C. R. and M. C. Anderson (2017), Estimating morning change in land surface temperature from
823 MODIS day/night observations: Applications for surface energy balance modeling, *Geophys. Res.*
824 *Lett.*, 44(19), 9723-9733, doi:10.1002/2017GL074952.

825 Haughton, N., G. Abramowitz, A. J. Pitman, D. Or, M. J. Best, H. R. Johnson, G. Balsamo, A. Boone, M.
826 Cuntz, B. Decharme, and P. A. Dirmeyer (2016), The plumbing of land surface models: Is poor
827 performance a result of methodology or data quality?, *J. Hydrometeor.*, 17(6), pp.1705-1723,
828 doi:10.1175/JHM-D-15-0171.1.

829 Holmes, T. R. H., C. R. Hain, W. T. Crow, M. C. Anderson, and W. P. Kustas (2018), Microwave
830 implementation of two-source energy balance approach for estimating evapotranspiration, *Hydrol.*
831 *Earth Syst. Sci.*, 22(2), 1351-1369, doi:10.5194/hess-22-1351-2018.

832 Holmes, T. R. H., W. T. Crow, C. R. Hain, M. Anderson, and W. P. Kustas (2015), Diurnal temperature
833 cycle as observed by thermal infrared and microwave radiometers, *Remote Sens. Environ.*, 158,
834 110–125, doi:10.1016/j.rse.2014.10.031.

835 Huber, P. J. (2011), Robust statistics, In *International Encyclopedia of Statistical Science*, pp. 1248-
836 1251, Springer.

837 Jung, M., M. Reichstein, P. Ciais, S. I. Seneviratne, J. Sheffield, M. L. Goulden, G. Bonan, A. Cescatti, J.
838 Chen, R. De Jeu, et al. (2010), Recent decline in the global land evapotranspiration trend due to
839 limited moisture supply, *Nature*, 467, 951-954, doi:10.1038/nature09396.

840 Kerr, Y. H., P. Waldteufel, J. P. Wigneron, S. Delwart, F. Cabot, J. Boutin, M. J. Escorihuela, J. Font, N.
841 Reul, C. Gruhier, and S. E. Juglea (2010), The SMOS mission: New tool for monitoring key elements
842 of the global water cycle, *Proc. IEEE*, 98(5), 666-687, doi:10.1109/JPROC.2010.2043032.

843 Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar (2000), A catchment-based
844 approach to modeling land surface processes in a general circulation model: 1. Model structure, *J.*
845 *Geophys. Res: Atmos.*, 105(D20), 24809-24822, doi:10.1029/2000JD900327.

846 Koster, R. D., P. A. Dirmeyer, Z. Guo, G. Bonan, E. Chan, P. Cox, C. T. Gordon, S. Kanae, E. Kowalczyk,
847 D. Lawrence, et al. (2004), Regions of strong coupling between soil moisture and precipitation,
848 *Science*, 305(5687), 1138-1140, doi:10.1126/science.1100217.

849 Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel (2006), World map of the Köppen-Geiger
850 climate classification updated. *Meteorologische Zeitschrift*, 15(3), 259-263, doi:10.1127/0941-
851 2948/2006/0130.

852 Kumar, S. V., P. A. Dirmeyer, C. D. Peters-Lidard, R. Bindlish, and J. Bolten (2018), Information
853 theoretic evaluation of satellite soil moisture retrievals, *Remote Sens. Environ.*, 204, 392-400,
854 doi:10.1016/j.rse.2017.10.016.

855 Li, F., W. P. Kustas, M. C. Anderson, T. J. Jackson, R. Bindlish, and J. H. Prueger (2006), Comparing the
856 utility of microwave and thermal remote-sensing constraints in two-source modeling over an
857 agricultural landscape, *Remote Sens. Environ.*, 101, 315-329, doi:10.1016/j.rse.2006.01.001.

858 Liang, X., D. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hydrologically based model of
859 land surface water and energy fluxes for GSMs, *J. Geophys. Res.*, 99(D7), 14,415–14,428,
860 doi:10.1029/94JD00483.

861 Loew, A., W. Bell, L. Brocca, C. E. Bulgin, J. Burdanowitz, X. Calbet, R. V. Donner, D. Ghent, A. Gruber,
862 and T. Kaminski (2017), Validation practices for satellite based earth observation data across
863 communities, *Rev. Geophys.*, 55(3), doi:10.1002/2017RG000562.

864 Liu, Y. Y., W. A. Dorigo, R. M. Parinussa, R. A. de Jeu, W. Wagner, M. F. McCabe, J. P. Evans, and A. I. J.
865 N. van Dijk (2012), Trend-preserving blending of passive and active microwave soil moisture
866 retrievals, *Remote Sens. Environ.*, 123, 280-297, doi:10.1016/j.rse.2012.03.014.

867 McColl, K. A., J. Vogelzang, A. G. Konings, D. Entekhabi, M. Piles, and A. Stoffelen (2014), Extended
868 triple collocation: Estimating errors and correlation coefficients with respect to an unknown target,
869 *Geophys. Res. Lett.*, 41(17), 6229-6236, doi:10.1002/2014GL061322.

870 Naeimi, V., K. Scipal, Z. Bartalis, S. Hasenauer, and W. Wagner (2009), An improved soil moisture
871 retrieval algorithm for ERS and METOP scatterometer observations, *IEEE Trans. Geosci. Remote*
872 *Sens.*, 47(7), 1999-2013, doi:10.1109/TGRS.2008.2011617.

873 Niu, G. Y., Z. L. Yang, K. E. Mitchell, F. Chen, M. B. Ek, M. Barlage, A. Kumar, K. Manning, D. Niyogi, E.
874 Rosero, M. Tewari, and Y. Xia (2011), The community Noah land surface model with
875 multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale
876 measurements, *J. Geophys. Res: Atmos.*, 116(D12), doi:10.1029/2010JD015139.

877 Njoku, E. G., T. J. Jackson, V. Lakshmi, T. K. Chan, and S. V. Nghiem (2003), Soil moisture retrieval
878 from AMSR-E, *IEEE Trans. Geosci. Remote Sens.*, 41(2), 215-229, doi:10.1109/TGRS.2002.808243.

879 Owe M., R. de Jeu, and T. Holmes (2008), Multisensor historical climatology of satellite-derived
880 global land surface moisture, *J. Geophys. Res: Earth Surf.*, 113(F1), doi:10.1029/2007JF000769.

881 Parinussa, R. M., T. R. H. Holmes, and R. A. M. de Jeu (2012), Soil moisture retrievals from the
882 WindSat spaceborne polarimetric microwave radiometer, *IEEE Trans. Geosci. Remote Sens.*, 50(7),
883 2683-2694, doi:10.1109/TGRS.2011.2174643.

884 Parinussa, R. M., T. R. Holmes, N. Wanders, W. A. Dorigo, and R. A. M. de Jeu (2015), A preliminary
885 study toward consistent soil moisture from AMSR2, *J. Hydrometeor.*, 16(2), 932-947,
886 doi:10.1175/JHM-D-13-0200.1.

887 Phillips, T. J. and S. A. Klein (2014), Land-atmosphere coupling manifested in warm-season
888 observations on the U.S. southern great plains, *J. Geophys. Res.: Atmos.*, 119(2), 509-528,
889 doi:10.1002/2013JD020492.

890 Rodell, M., P. R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C. Meng, K. Arsenault, B. Cosgrove, J.
891 Radakovich, M. Bosilovich, J. K. Entin, J. P. Walker, D. Lohmann, and D. Toll (2004), The global land
892 data assimilation system, *Bull. Amer. Meteor. Soc.*, 85(3), 381-394, doi:10.1175/BAMS-85-3-381.

893 Saha, S., S. Moorthi, H.-L. Pan, X. Wu, J. Wang, S. Nadiga, P. Tripp, R. Kistler, J. Woollen, D. Behringer,
894 et al. (2010), The NCEP climate forecast system reanalysis, *Bull. Amer. Meteor. Soc.*, 91(8), 1015-
895 1058, doi:10.1175/2010BAMS3001.1.

896 Saha, S., S. Moorthi, X. Wu, J. Wang, S. Nadiga, P. Tripp, D. Behringer, Y. T. Hou, H.-y. Chuang, M.
897 Iredell, M. Ek, J. Meng, R. Yang, M.P. Mendez, H. van den Dool, Q. Zhang, W. Wang, M. Chen, and E.
898 Becker (2014), The NCEP climate forecast system version 2, *J. Clim.*, 27(6), 2185-2208,
899 doi:10.1175/JCLI-D-12-00823.1.

900 Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling
901 (2010), Investigating soil moisture-climate interactions in a changing climate: A review, *Earth Sci.*
902 *Rev.*, 99(3), 125-161, doi:10.1016/j.earscirev.2010.02.004.

903 Seneviratne, S. I., M. Wilhelm, T. Stanelle, B. Hurk, S. Hagemann, A. Berg, F. Cheruy, M. E. Higgins, A.
 904 Meier, V. Brovkin, et al. (2013), Impact of soil moisture-climate feedbacks on CMIP5 projections:
 905 First results from the GLACE-CMIP5 experiment, *Geophys. Res. Lett.*, 40(19), 5212-5217,
 906 doi:10.1002/grl.50956.

907 Sörensson, A. A., and R. C. Ruscica (2018), Intercomparison and Uncertainty Assessment of Nine
 908 Evapotranspiration Estimates Over South America, *Water Resour. Res.*, 54(4), 2891-2908,
 909 doi:10.1002/2017WR021682.

910 Stoffelen, A. (1998), Toward the true near-surface wind speed: Error modeling and calibration
 911 using triple collocation, *J. Geophys. Res.: Oceans* (1978–2012), 103(C4), 7755-7766,
 912 doi:10.1029/97JC03180.

913 Su, C. H., D. Ryu, W. T. Crow, and A. W. Western (2014), Beyond triple collocation: Applications to
 914 soil moisture monitoring, *J. Geophys. Res.: Atmos.*, 119(11), 6419-6439,
 915 doi:10.1002/2013JD021043.

916 Teuling, A. J., R. Uijlenhoet, and P. A. Troch (2005), On bimodality in warm season soil moisture
 917 observations, *Geophys. Res. Lett.*, 32, L13402, doi:10.1029/2005GL023223.

918 Tuttle, S. and G. Salvucci (2016), Empirical evidence of contrasting soil moisture-precipitation
 919 feedbacks across the United States, *Science*, 352(6287), 825-828, doi:10.1126/science.aaa7185.

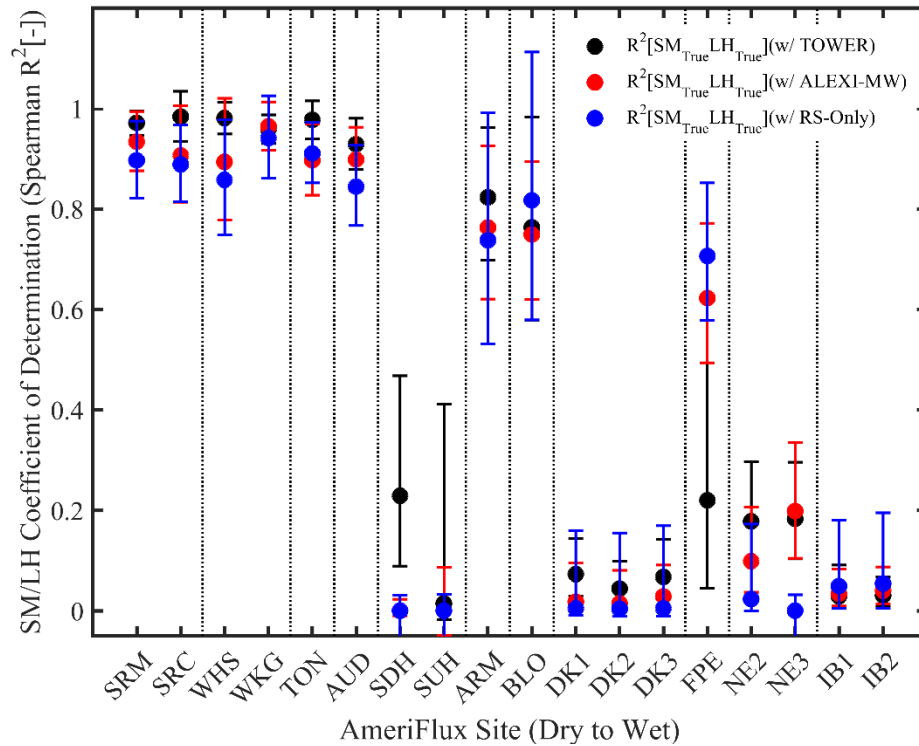
920 Ukkola, A., M. De Kauwe, A. Pitman, M. Best, G. Abramowitz, V. Haverd, M. Decker, and N. Haughton
 921 (2016), Land surface models systematically overestimate the intensity, duration and magnitude of
 922 seasonal-scale evaporative droughts, *Environ. Res. Lett.*, 11(10), 104012.

923 Ukkola, A. M., A. J. Pitman, and M. G. De Kauwe, G. Abramowitz, J. Kala, and Y-P. Wang (2016),
 924 Modelling evapotranspiration during precipitation deficits: identifying critical processes in a land
 925 surface model, *Hydrol. Earth Syst. Sci.*, 20(6), 2403-2419, doi:10.5194/hess-20-2403-2016.

926 Van den Hurk, B., M. Best, P. Dirmeyer, A. Pitman, J. Polcher, and J. Santanello (2011), Acceleration
 927 of land surface model development over a decade of GLASS, *Bull. Amer. Meteor. Soc.*, 92(12), 1593-
 928 1600, doi:10.1175/BAMS-D-11-00007.1.

929 Van den Hurk, B., H. Kim, G. Krinner, S. I. Seneviratne, C. Derksen, T. Oki, H. Douville, J. Colin, A.
 930 Ducharne, F. Cheruy, N. Viovy, et al. (2016), LS3MIP (v1.0) contribution to CMIP6: the Land Surface,
 931 Snow and Soil moisture Model Intercomparison Project-aims, setup and expected outcome, *Geosci
 932 Model Dev.*, 9, 2809-2832, doi:10.5194/gmd-9-2809-2016.

933 Wielicki, B. A., B. R. Barkstrom, E. F. Harrison, R. B. Lee III, G. L. Smith, and J. E. Cooper (1996),
 934 Clouds and the Earth's Radiant Energy System (CERES): An earth observing system
 935 experiment, *Bull. Amer. Meteor. Soc.*, 77(5), 853-868, doi:10.1175/1520-0477(1996)077<0853:
 936 CATERE>2.0.CO;2.



939

940

941 Figure 1. TC-based SM/LH coupling strength estimates acquired from a variety of sources
 942 over a set of AmeriFlux sites (see Appendix B). Black symbols (“TOWER”) replicate earlier
 943 results in Crow et al. (2015) based on using ground/tower-based SM and LH observations
 944 (plus Noah SM, Noah LH, PMW SM, and ALEXI-TIR LH) to complete the triplet. Red symbols
 945 (“ALEXI-MW”) capture the impact of using ALEXI-MW LH in place of flux tower observations.
 946 Blue symbols (“RS-Only”) capture the impact of further replacing ground-based SM
 947 observations with AMW SM retrievals (thus eliminating all dependence on ground-based
 948 observations). The consistency demonstrates the applicability of selected RS-based SM and
 949 LH in place of ground-based data for TC analysis.

950

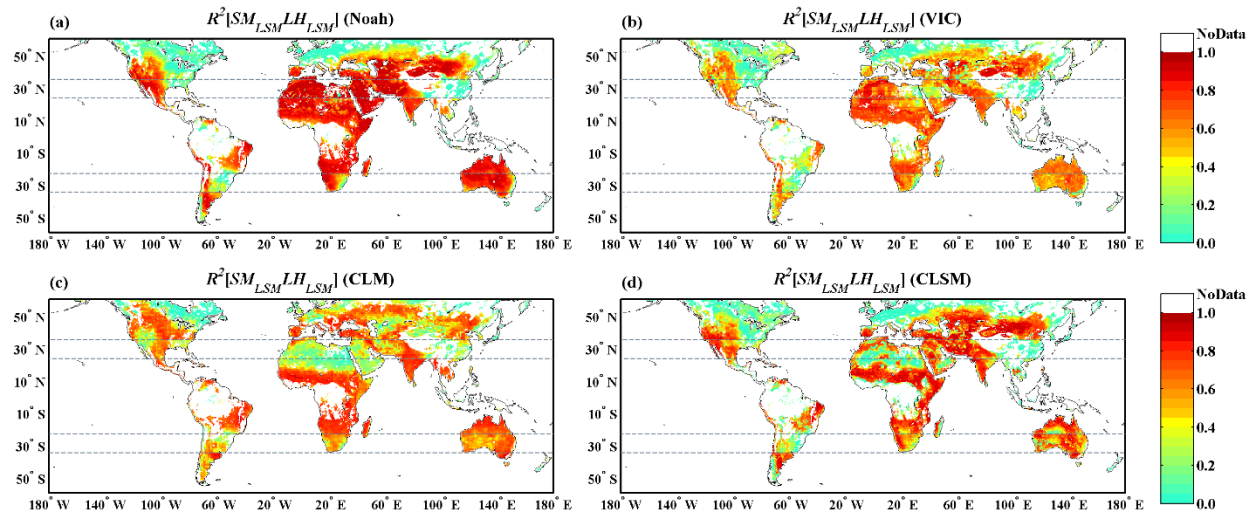


Figure 2. Variations of predicted warm-season SM/LH coupling strength between the LSMs: (a) Noah, (b) VIC, (c) CLM, and (d) CLSM. Parallel lines at different latitudes indicate the delineation of tropical, sub-tropical and temperate zones for defining the warm-season period. White land grid cells indicate regions with insufficient RS-based retrievals for comparison ("NoData" in legend).

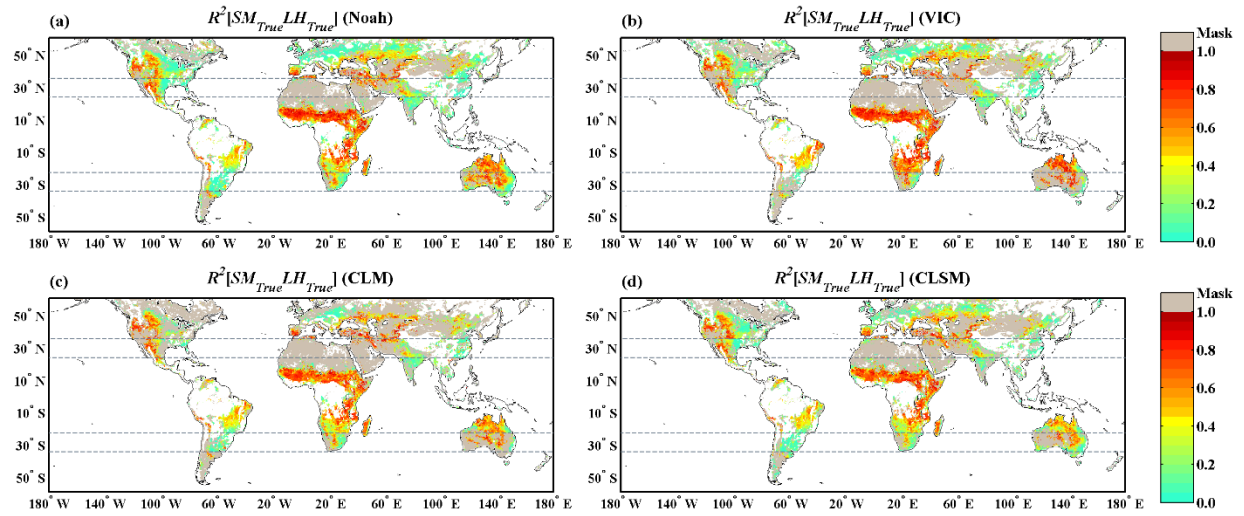


Figure 3. TC-based estimates of warm-season SM/LH coupling strength with LSM SM and LH datasets obtained from the (a) Noah, (b) VIC, (c) CLM, and (d) CLSM LSMs. The generally consistent magnitudes and agreement in spatial patterns demonstrate the robustness of TC results with respect to LSM model choice.

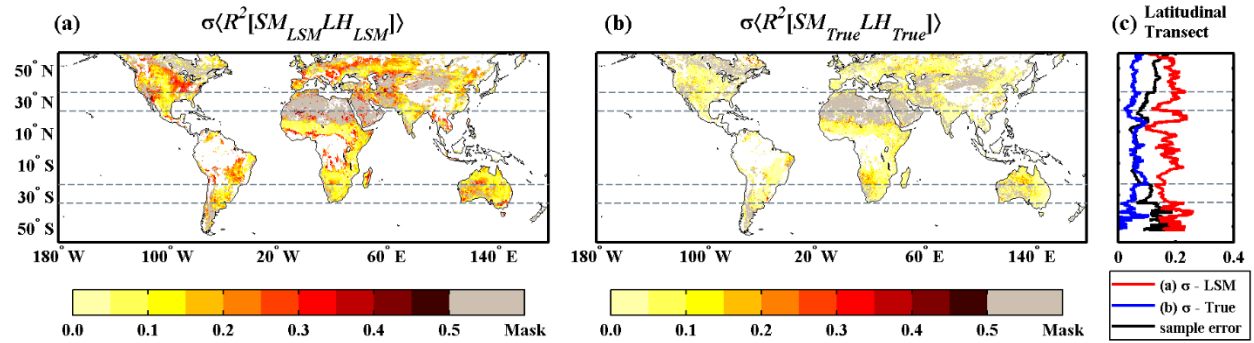


Figure 4. Inter-model variations (expressed as standard deviations) in warm-season SM/LH coupling strength across different (a) LSMs and (b) TC-based estimates with independent LSM as input. Mean standard deviations along the latitudinal transect are shown in (c), as well as the 1-sigma sampling error (black line) obtained from a boot-strapping analysis.

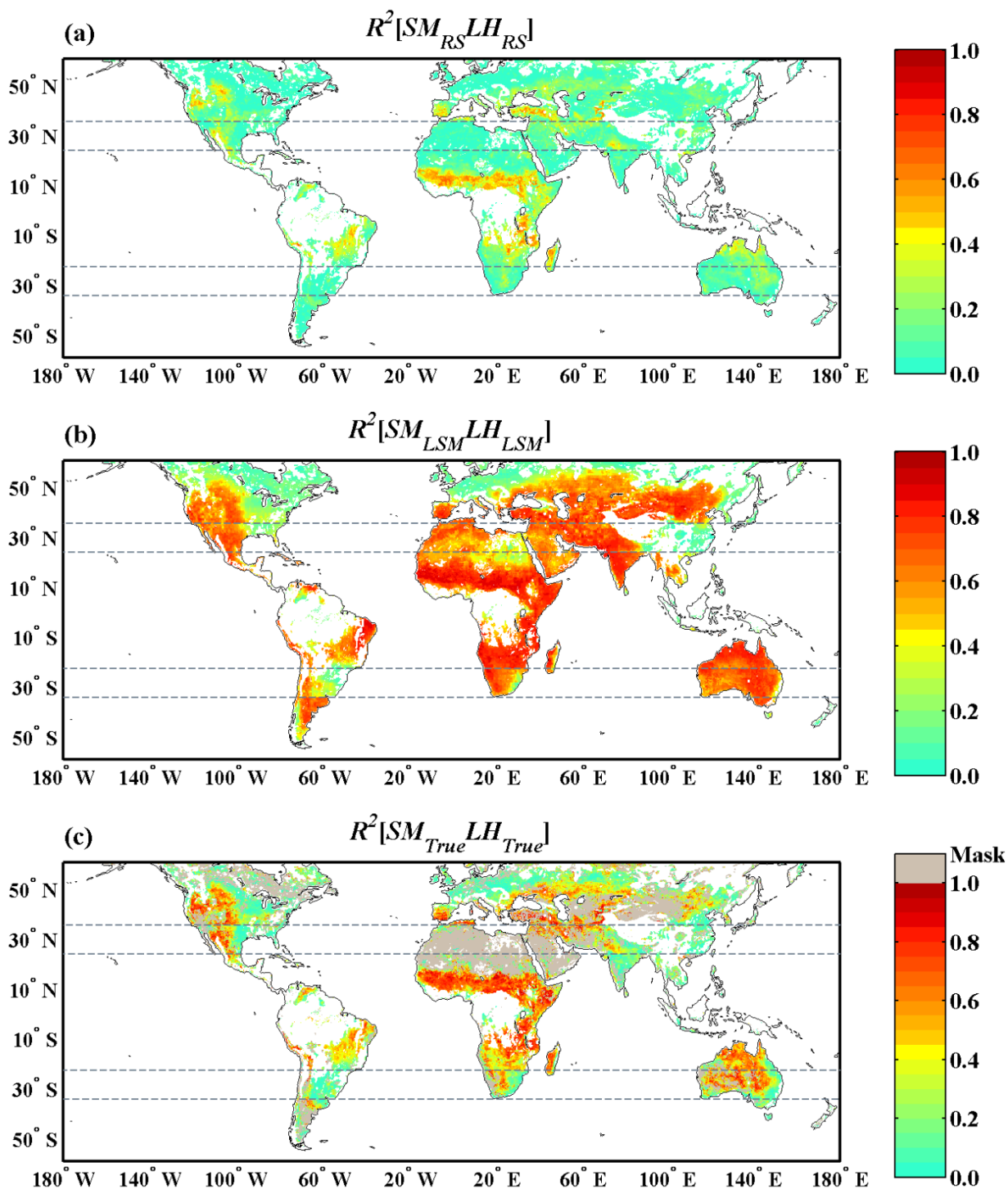


Figure 5. Global mean estimates of warm-season SM/LH coupling derived from (a) RS-based products, (b) LSMs, and (c) TC-based estimates.

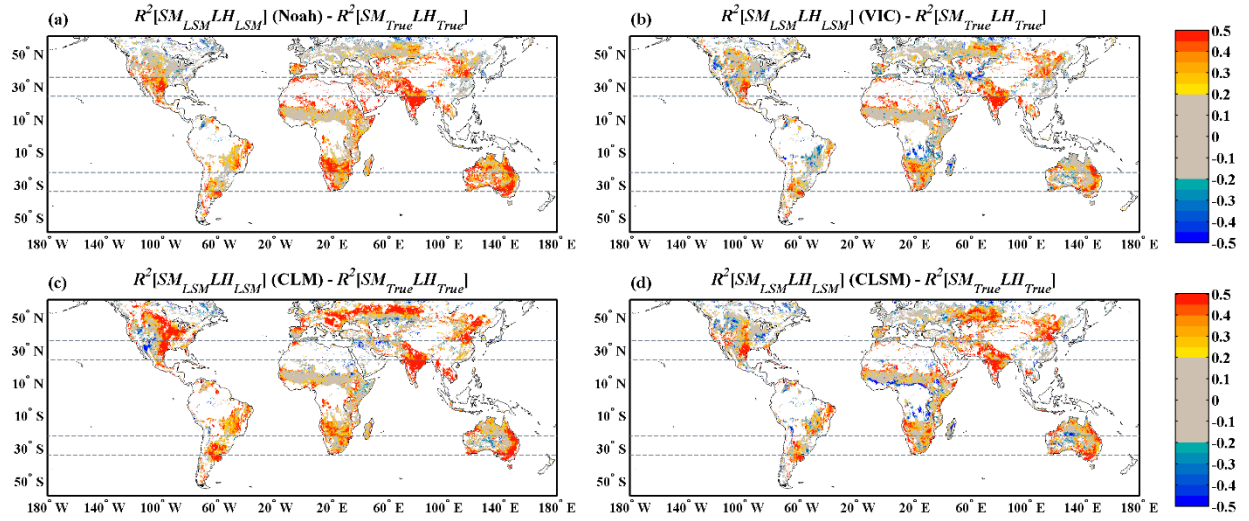
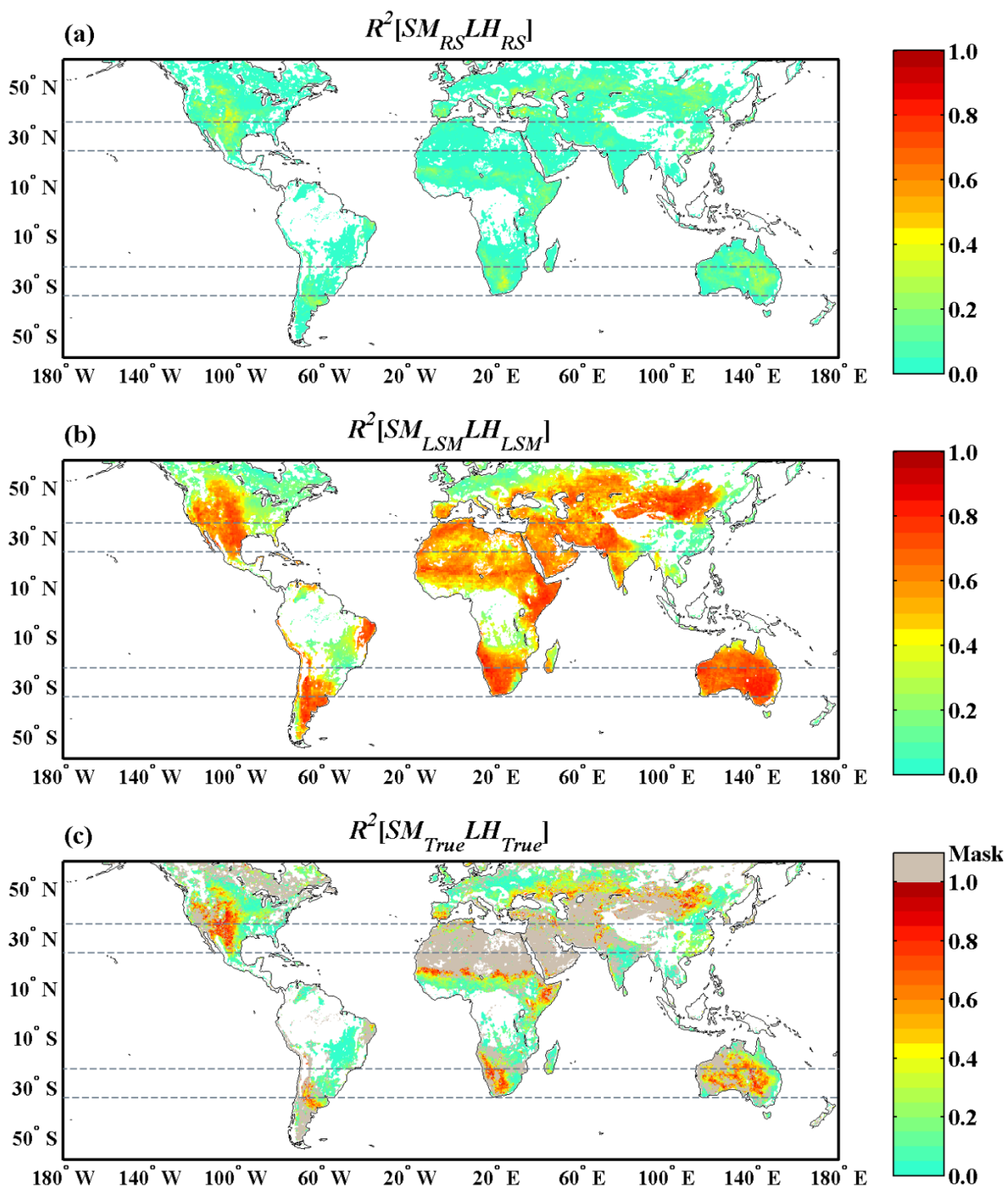


Figure 6. Differences of warm-season SM/LH coupling strength between model predictions and TC-based estimates for the (a) Noah, (b) VIC, (c) CLM, and (d) CLSM LSMs. Grids within the 2-sigma confidence interval are masked (see grey shading).



992

993

994 Figure 7. Same as Figure 5 but for SM and LH anomaly time series.

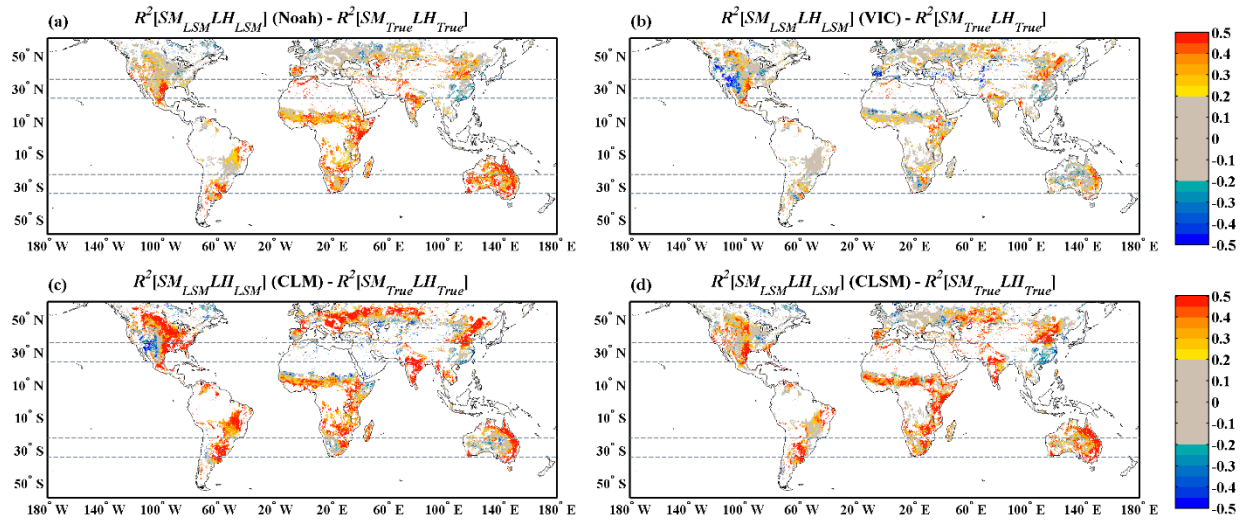


Figure 8. Same as Figure 6 but for SM and LH anomaly time series.

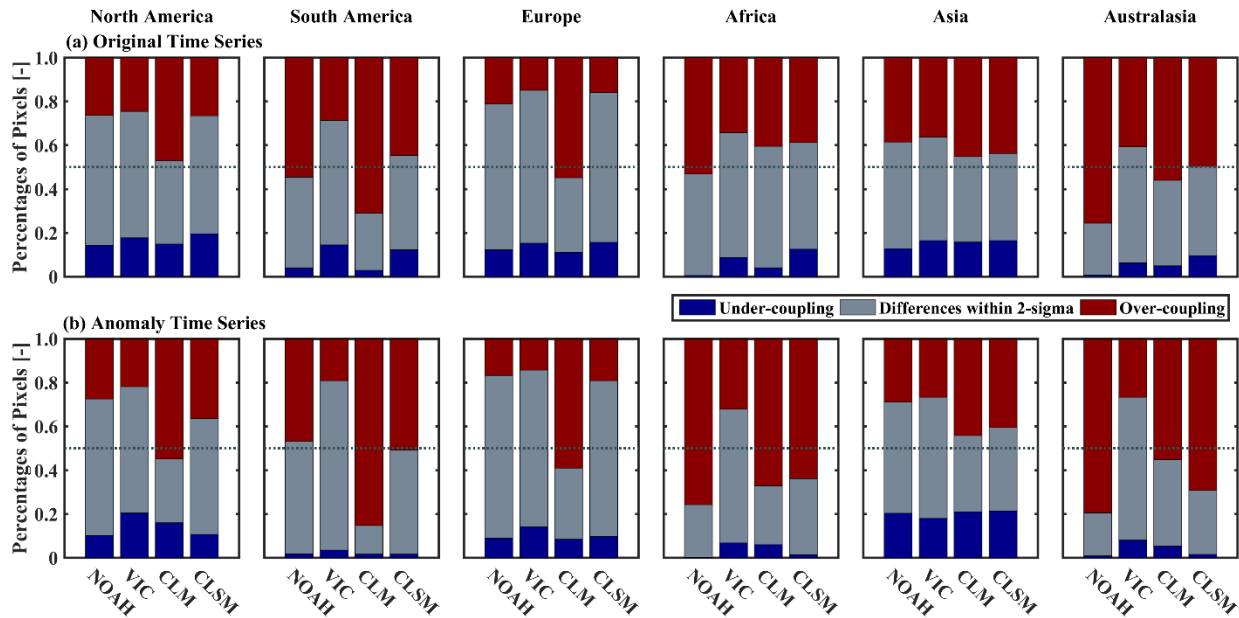


Figure 9. Percentages of pixels with SM/LH over-coupling (dark red) and under-coupling (dark blue) in LSMs' predictions as compared to the TC-based estimates for six continental regions. Results are based on both (a) original and (b) anomaly time series of SM and LH. Percentages of pixels with (non-significant) differences between LSM- and TC-based coupling strength estimates within a 2-sigma confidence interval are in gray.

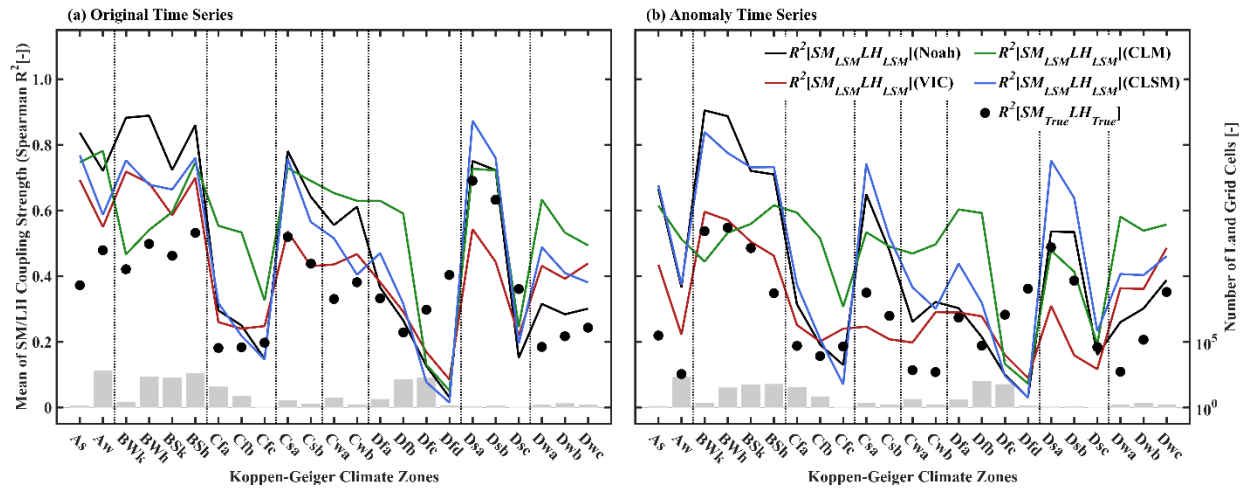


Figure 10. Aggregated mean values of warm-season SM/LH coupling strength for various Köppen-Geiger climate zones based on both (a) original and (b) anomaly time series analysis. Grey bars (and right vertical axis) indicate the number of grids with valid TC results.

Table

Table 1. Basic information for SM and LH used in the analysis.

| Product Name | Key Sources | Spatial Resolution | Temporal Resolution | Temporal Range |
|--------------------------|--|--------------------|---------------------|----------------|
| PMW SM (ESA CCI PASSIVE) | AMSR-E LPRM SM AMSR-2 LPRM SM SMOS LPRM SM | 0.25° | Daily | 1978–2015 |
| AMW SM (EUMETSAT) | ASCAT TU-Wien SM | 25 km | Daily | 2007 onwards |
| ALEXI-TIR | MODIS LST | 0.05° | Weekly | 2003–2013 |
| ALEXI-MW | MW-LST (Section 2.3) | 0.25° | Weekly | 2003–2013 |
| GLDAS-2.1 NOAH | GDAS, GPCP, AGRMET (Section 2.4) | 0.25° | 3-hourly | 2000 onwards |
| GLDAS-1 VIC | GDAS, GPCP, AGRMET | 1.0° | 3-hourly | 1979 onwards |
| GLDAS-1 CLM | GDAS, GPCP, AGRMET | 1.0° | 3-hourly | 1979 onwards |
| GLDAS-2.0 CLSM | Princeton Meteorological Forcing | 0.25° | Daily | 1948-2014 |

Table 2. Explanations for climate classification index (adopted from Kottek et al. (2006)).

| Main Climate | Precipitation | Temperature |
|--------------------------|-----------------------|---------------------------------|
| A: equatorial | W: Desert | h: hot arid |
| B: arid | S: Steppe | k: cold arid |
| C: warm temperate | f: fully humid | a: hot summer |
| D: continental | s: summer dry | b: warm summer |
| E: polar | w: winter dry | c: cool summer |
| | m: monsoonal | d: extremely continental |
| | | F polar frost |
| | | T polar tundra |